

# ajer

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXIII    NUMBER 1    MARCH 1977

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA    ●    EDMONTON



## THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

.....

Published quarterly in March, June, September and December  
by the  
Faculty of Education  
of  
The University of Alberta

.....

### Faculty Publications Committee

*K. A. Neufeld*, Department of Elementary Education  
*J. Chambers*, Department of Educational Psychology  
*A. K. Clark*, Department of Industrial and Vocational Education  
*A. G. Konrad*, Department of Educational Administration  
*M. R. Lupul*, Department of Educational Foundations  
*R. G. Martin*, Department of Secondary Education  
*C. H. Preitz*, Department of Industrial and Vocational Education

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

Chairman: *K. A. Neufeld*

Editor: *Andrew K. Clark*

Secretary: *Mrs. A. Onishenko*

.....

AJER gratefully acknowledges support from Canada Council and AACES. AJER is indexed in the *Canadian Education Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, and *Sociology of Education Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available in bound and unbound volumes; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, The University of Alberta, Edmonton, Alberta, Canada T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



# The Alberta Journal of Educational Research

Volume XXIII, No. 1


March, 1977

## CONTENTS

Effects of Age and Sex on the Acquisition of Form Discrimination .....	1
<i>J. Tyler and R. C. Hardy</i>	
Linguistic Predictors of Properties of Set .....	11
<i>H. L. Janzen and H. J. Hallworth</i>	
Children's Behavior in Solving Spatial Problems .....	22
<i>L. D. Nelson and T. E. Kieren</i>	
The Contributions of L. S. Vygotsky to Cognitive Psychology .....	31
<i>S. Phillips</i>	
The Developmental Sequence of Cognition as Revealed by Wrong Answers .....	43
<i>J. C. Powell</i>	
The Use of Infra-red False Color Satellite Images by Grades 3, 4, and 5 Pupils and Teachers .....	52
<i>J. Kirman</i>	
Teacher Self-Actualization and Pupil Control Ideology-Behavior Consistency .....	65
<i>R. L. Noll, D. J. Willower and J. J. Barnette</i>	
The Interactive Relationships of Teacher Directiveness and Student Authoritarianism and Dogmatism to Grades and Satisfaction .....	71
<i>S. J. H. McCann and R. J. Fisher</i>	
Consultative Practices of Elementary School Teachers: Some Research Findings .....	85
<i>M. Haughey, E. A. Holdaway, and J. M. Small</i>	

FACULTY OF EDUCATION  
*The University of Alberta*





Digitized by the Internet Archive  
in 2024 with funding from  
University of Alberta Library

[https://archive.org/details/ajer\\_1977](https://archive.org/details/ajer_1977)



JOANNA TYLER

and

R. C. HARDY

University of Maryland

## Effects of Age and Sex on the Acquisition of Form Discrimination

*Forty subjects aged three (10 males and 10 females) and four (10 males and 10 females) were presented with a rotated, letter-like form discrimination matching task. The purpose was to investigate developmental changes across age and sex in acquisition of form discrimination. Results indicated that: (1) horizontal-vertical rotations were discriminated better than oblique rotations; (2) four-year-old females performed better than four-year-old males on discriminating oblique rotations; and (3) four-year-olds performed better than three-year-olds on both kinds of rotations. Results suggested that age, sex, and type of rotation are important when training for form discrimination ability. (Ms. Tyler and Mr. Hardy are with the Institute for Child Study at the University of Maryland.)*

The age at which children acquire the ability to distinguish variants of one visual pattern from variants of another has been of interest to investigators concerned with the process of how children learn to read. Gibson (1965) postulated that there are three phases in learning to read: (1) differentiating graphic symbols, (2) decoding letters to sounds, and (3) using progressively higher-order units of structure. This study focuses on the first phase of this process.

Two theories, maturational and learning, have been used to define the process by which a child acquires the ability to discriminate form. Two interactionists, Piaget and Inhelder (1956), suggested support for a maturational explanation of this phenomenon. Gibson (1965) provided support for a learning point of view. Gibson postulated a distinctive features hypothesis suggesting that form discrimination is a result of learning to detect the invariants of form. While this hypothesis has received some



support, attempts to use letter-like forms to train children's form discrimination ability in order that it might be transferred to new forms has only been successful with five-year-old subjects (Silver & Rollins, 1973). Silver and Rollins also noted that five-year-old males perform better than five-year-old females. It appears that learning to recognize the distinctive features of form is, perhaps, associated with readiness level of the child and may be linked to a sex variable. Consistent with this, Williams (1969) investigated the effectiveness of different training methods on acquisition of distinctive features and found that successful readiness training depends on the cognitive level of the child and the kinds of rotational transformation material presented. Kindergarten children made more discrimination errors on left-right orientations of letter-like forms than they did on up-down orientations of these forms, although discrimination training was more successful on the whole than reproduction training. Children with more kindergarten experience were more successful in discrimination training tasks than beginning kindergarten children. From these studies, it appears that the important variables are maturational level, sex, and type of transformation.

Despite this general conclusion in the literature, there is considerable research designed to teach children form discrimination techniques regardless of their readiness level. The work of Gibson, Gibson, Pick and Osser (1962) represents one of the few attempts to investigate the developmental aspects of form discrimination learning across age (four to eight) and sex. Results indicated that form discrimination ability increases from ages four to eight and that females aged four and five outperform males of similar age. Other research supports the notion that children's ability to discriminate and recognize different symbols or pattern of line figures begins to develop at age four and continues to improve until approximately age eight (Davidson, 1934, 1935; Welch, 1939; Ghent, 1961; Gibson et al., 1962; Rudel & Teuber, 1963; Bishop, 1964; and Rosenberg & Birch, 1969).

The differing performance among sex and age groups on various form discrimination tasks may have been a function of methodological variation. The kinds of form stimuli used to investigate this phenomenon appear to effect children's ability to discriminate the stimuli form. Jeffrey (1958) used left-right orientations of stick figures with four-year-old children to investigate the kind and amount of experience necessary in learning discrimination tasks. He found that four-year-old males and females could discriminate left-right orientations with ease. Jeffrey concluded that the critical time to stimulate form discrimination learning is at age four. While Williams (1969) found that 4½- to 5½-year-old children discriminate up-down orientations more easily than left-right orientations, she neither extended her investigation to include oblique orientations nor did she note sex differences.

Rudel and Teuber (1963) found that children aged three, four, and five failed to discriminate oblique orientations of a line. Both three- and four-year-olds could discriminate on the vertical and horizontal orientation; three-year-olds made more errors than four-year-olds. This ability to discriminate may have been due to the simplicity of the stimuli (one vertical or one horizontal line matched or mismatched with another line). Rudel and



Teuber's conclusion suggested that horizontal and vertical orientations are learned before the oblique orientation. More recent studies lend some support to this conclusion (Kolars & Perkins, 1969a, 1969b; Olson, 1970).

The previously cited research of Rudel and Teuber indicates that little difference exists between three- and four-year-old children on simple line discriminations. When the discriminations were horizontal and vertical, both ages could perform successfully but not significantly different from each other. When oblique orientations were used, neither group could perform discriminations successfully.

Jeffrey's research (1958) also supports this since he found that four-year-olds could discriminate left-right orientations of stick figures. It might be argued that certain forms, such as letter-like forms, are more easily assimilated by four-year-olds, due to their early contact with similar forms in preschool education, than three-year-olds who have not had as much contact with these forms. In addition, use of letter-like forms should make it easier for four-year-old children to discriminate the more difficult oblique orientations than when irregular patterns of lines are used as in the Rudel and Teuber study.

The purpose of the present investigation is to look at developmental changes in fine form discrimination ability between males and females aged three and four when letter-like forms are presented in horizontal-vertical and oblique rotations. It was predicted that children performed better when letter-like stimuli were rotated in the horizontal-vertical position rather than in the oblique position; that four-year-old females performed better than four-year-old males; and that four-year-olds (male and female) performed better than three-year-olds (male and female) on total, horizontal-vertical, and oblique letter-like form discriminations. From the above discussion, three null hypotheses are postulated:

1. For all children, there is no difference in performance on stimuli in the horizontal-vertical rotation than on stimuli in the oblique rotation.
2. There is no difference in the performance of four-year-old females from four-year-old males on total correct discriminations, horizontal-vertical discriminations, and oblique discriminations.
3. There is no difference in the performance of four-year-old males and females from three-year-old males and females in their ability to make correct discriminations on total correct discriminations, horizontal-vertical discriminations, and oblique discriminations.

### *Method*

#### *Subjects*

Twenty children aged approximately three and one-half (10 males and 10 females) and 20 children aged approximately four and one-half (10 males and 10 females) were randomly selected from a subject pool of 83 children from five suburban preschools in the Baltimore, Maryland, metropolitan area and randomly assigned within each age group to one of two treatment groups. Thus, there was a total of four groups with 10 subjects each based on age and sex.



Materials

There was one set of 32 stimuli cards. Each card was a 20- by 6-inch (50.80 cm × 15.24 cm) masonite board bearing five 4- by 4½-inch (10.16 cm × 11.43 cm) black and white, high gloss photographed letter-like forms. (The letter-like forms on each card were selected from forms generated by Gibson et al., 1962.) Sixteen of the 32 stimuli cards represented horizontal and vertical rotations while the remaining 16 represented oblique rotations. Of the 16 horizontal-vertical rotations of letter-like forms, 8 represented horizontal rotations and 8 represented vertical rotations. In the horizontal rotation, the target figures were rotated 90° to the right and 90° to the left while in the vertical rotation, the letter-like forms were rotated 180° or not at all. Of the 16 oblique rotations, 8 letter-like forms represented rotations of 35° (four forms were rotated to the left and four to the right) and 8 represented rotations of 145° (four forms were rotated to the left and four to the right). Examples of these stimuli rotations can be found in Figure 1.

FIGURE 1  
AN EXAMPLE OF THE HORIZONTAL-VERTICAL AND OBLIQUE ROTATIONS

TARGET		HORIZONTAL-VERTICAL ROTATIONS					
		90°R	90°L		360°U	180°D	
1.	<div>90° L</div>						
2.	<div>180° D</div>						
OBLIQUE ROTATIONS							
		35°R	145°R		35°L	145°L	
3.	<div>35° L</div>						
4.	<div>35° R</div>						



It was the subject's task to match the target figure, which was rotated in the different positions described above, to one of four identical letter-like forms representing different rotations, one of which matched the rotation of the target figure. Thus five different positions were needed on each card. Position 1 was located to the left of the subject. It always displayed the target figure which was to be matched with one of the other four letter-like forms displayed in positions 2 through 5. This placement encouraged subjects to search the stimuli from left to right, the direction appropriate for reading English. Letter-like stimuli in positions 2 through 5 were randomized to prevent subjects from developing strategies about the position of the correct match.

### *Procedure*

Each subject received one training trial. The following instructions were given:

You sit here and I'll sit next to you and show you how to play a game. Look at all of these pictures (point). Now look at this picture (point). Next, look at these pictures and choose one of them that looks just like this one (point). You put your finger on the picture that looks just like this one.

If the subject responded correctly on the demonstration trial, the experimenter proceeded to give the 32 stimuli cards of one trial each, saying each time, "Which picture looks just like this one?" If the subject did not respond correctly to the demonstration trial, the instructions were repeated. If the subject failed to understand the instructions on the second reading, he was dropped from the pool of subjects and another subject was randomly chosen from the pool. Six of the originally selected 40 subjects were dropped.

The 32 stimuli cards were randomly split into four groups of eight stimuli cards. Each group contained one instance of each of the various rotations (90° left and right, 35° left and right, 145° left and right, 180°, and no rotation). The order of presentation of cards within a group was also randomly assigned. Each subject did all four groups, but the order of presentation was randomly assigned to him. All possible permutations (24) were utilized before any permutation was duplicated. This procedure was done in an effort to prevent a facilitating effect of starting position.

### *Data Analysis*

To determine if any effect existed due to starting position, a  $2 \times 4$  factorial design with two levels of stimuli rotation (horizontal-vertical and oblique) and four levels of starting position was used. The dependent variables were the number of correct responses on the horizontal-vertical and the number of correct responses on the oblique rotations. The level of significance for rejection of the null hypotheses was  $p < .05$ .

To test hypothesis 1 and the assumption that starting block position had no effect on performance, a  $2 \times 4$  factorial analysis of variance was run between rotation (horizontal-vertical vs. oblique) and four levels of starting block position. To test hypotheses 2 and 3, a  $2 \times 2$  factorial design with two levels of age (three-year-old and four-year-old) and two levels of sex was used. The dependent variables were the number of correct responses made on the horizontal-vertical rotation, the number of correct responses made on



the oblique rotation, and the total number of correct responses made on all 32 trials.

### Results

#### Hypothesis 1

The means and standard deviations for starting block positions and rotations are shown in Table 1. The higher the score the better was the performance on the task. It can be seen from this table that the means for the horizontal-vertical rotations were higher than for the oblique rotations which was in the direction predicted.

TABLE 1  
MEANS AND STANDARD DEVIATIONS ON STARTING  
BLOCK POSITION AND ROTATION

	Block Position			
	1	2	3	4
Oblique Rotation	7.00 (2.36)	6.47 (3.52)	6.25 (3.09)	5.92 (2.43)
Horizontal-Vertical Rotation	10.33 (2.66)	13.53 (3.64)	10.75 (2.63)	9.23 (4.19)

Note. Standard deviations are in parentheses.

Table 2 presents a summary of the  $2 \times 4$  factorial analysis of variance. The results indicate that starting block position had no effect on performance,  $F(1, 72) = .49$ ,  $p > .05$ . Thus, the assumption that starting block position does not effect performance was supported by the data. Horizontal-vertical rotations were discriminated significantly better than oblique orientations,  $F(1, 72) = 25.17$ ,  $p < .001$ . Thus the first null hypothesis that there is no difference for all children in performance between horizontal-vertical rotations and oblique rotations was rejected.

#### Hypotheses 2 and 3

Hypotheses 2 and 3 were tested by performing three  $2 \times 2$  factorial analyses of variance. The means and standard deviations for age and sex on the total trials, the 16 horizontal-vertical trials, and the 16 oblique trials are shown in Table 3. The means show that male and female four-year-olds performed equally well on the horizontal-vertical trials, but females had a higher mean performance than males on both total trials and oblique trials. Three-year-old males and females performed about equally well on all the dependent measures. Male and female four-year-olds performed better than male and female three-year-olds on all the dependent measures. Summaries of the analyses of variance can be found in Table 4.

Results of the analyses of variance for main effect of sex indicate that females performed significantly better than males on the oblique dimension,



TABLE 2  
FACTORIAL ANALYSIS OF VARIANCE BETWEEN ROTATION  
AND STARTING BLOCK POSITION

Source	SS	df	MS	F
Rotation	281.25	1	281.25	25.17*
Block Position	16.44	3	5.48	.49
Rotation x Block	3.73	3	1.24	.11
Within	804.54	72		
Total	1105.95	79		

\*  $p < .001$

TABLE 3  
MEANS AND STANDARD DEVIATIONS FOR AGE AND SEX  
ON EACH CRITERION

	Age 4		Age 3	
	Male	Female	Male	Female
Horizontal-Vertical Rotation	11.21 (3.42)	11.70 (3.33)	8.81 (3.01)	8.70 (3.83)
Oblique	5.60 (1.89)	8.90 (2.84)	4.90 (2.33)	6.00 (3.05)
Total	16.90 (4.77)	20.60 (5.21)	13.70 (4.66)	14.70 (6.39)

Note. Standard deviations are in parentheses.

$F(1, 36) = 7.31, p < .01$ . There were no significant results for the horizontal-vertical,  $F(1, 36) = .03, p < .05$ , or total responses,  $F(1, 36) = 1.96, p > .05$ . A Neuman-Keuls procedure performed on the oblique measure indicated that four-year-old females did better than four-year-old males, three-year-old males, and three-year-old females at  $p < .05$  level of significance. Since there were no main effects for sex on the horizontal-vertical or total response measures, a Neuman-Keuls procedure was not performed on these data. Therefore, results of this study only partially supported the predicted position. The second null hypothesis that there was no difference in the performance of four-year-old females from four-year-old males on total correct discriminations, horizontal-vertical discriminations, and oblique discriminations was rejected for only the oblique measure.

Results of the analyses of variance for the main effect of age indicated that four-year-old children performed significantly better than three-year-old children on total responses,  $F(1, 36) = 7.35, p < .01$ , horizontal-vertical responses,  $F(1, 36) = 6.26, p < .01$ , and oblique responses,  $F(1, 36) = 4.90,$

TABLE 4  
FACTORIAL ANALYSES OF VARIANCE FOR TOTAL CORRECT RESPONSES

Source	SS	df	MS	F
(a) <u>Total Correct Responses</u>				
Age	207.03	1	207.03	7.35**
Sex	55.23	1	55.23	1.96
Age x Sex	18.23	1	18.23	.64
Within	1013.50	36	28.15	
Total	1293.98	39		
(b) <u>Horizontal-Vertical Responses</u>				
Age	72.90	1	72.90	6.26**
Sex	.40	1	.40	.03
Age x Sex	.90	1	.90	.07
Within	419.40	36	11.65	
Total	493.60	39		
(c) <u>Oblique Responses</u>				
Age	32.40	1	32.40	4.90*
Sex	48.40	1	48.40	7.31**
Age x Sex	12.10	1	12.10	1.82
Within	238.20	36	6.61	
Total	331.10	39		

\*  $p < .05$

\*\*  $p < .01$

$p < .05$ . These results supported the predicted position. Therefore, the third null hypothesis that no difference in performance exists for four-year-old males and females from three-year-old males and females in their ability to make correct discriminations on total trials, horizontal-vertical trials, and oblique trials was rejected.

Discussion

Since all children perform better on stimuli in the horizontal-vertical rotation than on stimuli in the oblique rotation, the first null hypothesis is rejected. This, therefore, supports the previously mentioned positions of Rudel and Teuber (1963) and Jeffrey (1958). Furthermore, since there was no difference in starting block position, randomization of starting positions served as an adequate control in measuring the effect of rotation.

Although four-year-old females perform better on oblique rotations than do four-year-old males, they did not perform better on horizontal-vertical or total rotations. Consequently, the second null hypothesis which states that no difference in performance between four-year-old females and males would exist on total, horizontal-vertical, and oblique rotations was rejected for only



oblique rotations. In addition, a significant mean comparison indicates that four-year-old females perform better than three-year-old females on the oblique rotation. Three-year-old males and females perform similarly on the oblique orientation. While it appears that oblique form discrimination tends to crystalize in females during their fourth year, this does not appear to be true for males.

There are no differences between four-year-old males and females in discriminating letter-like forms in horizontal-vertical rotations. Although Rudel and Teuber (1963) and Olson (1970) reported that the oblique rotations are more difficult than horizontal-vertical rotations for young children to discriminate, results of this study suggest that this relationship is true for all young males but not for four-year-old females.

An explanation for the differential ability between males and females to discriminate form is possibly a function of different neurological development. Dick (1976) mentioned that Held supports the notion that there are two visual systems, one specialized for form perception and another specialized for space perception. Males and females have been shown to differ in their ability to use these two visual systems. For instance, females are more adept at two-dimensional kinds of perception, while males tend to perform better on spatial orientations where use of a three-dimensional system is most effective. If males and females initially respond in neurologically different ways, then four-year-old females' ability to excel four-year-old males in discriminating the rather difficult oblique angle appears to demonstrate support for a maturational theory of form discrimination ability. This, coupled with the notion that four-year-olds have more experience with form than three-year-olds, may suggest that females benefit more than males from their initial letter experiences as a result of their differential neurological development.

Since four-year-old children perform better than three-year-old children on all form discrimination criteria, the third null hypothesis that no difference exists in performance of four-year-old males and females from three-year-old males and females in their ability to make correct discriminations on total trials, horizontal-vertical trials, and oblique trials was rejected. It is suggested that this difference might be due not only to maturational differences, but also to the early learning experiences of the child, particularly in the preschool setting. The four-year-old children in this study had one more year of school experience than did the three-year-old children. Perhaps the task of letter-like forms was more familiar to the four-year-old group than it was to the younger children. The significant difference between age groups may, however, only be a result of the high scores achieved by the female subjects in the four-year-old group.

The suggestion that females use their neurological structure for analyzing two-dimensional form at an earlier age than males tends to support a maturational theory for form discrimination ability. This, in conjunction with the learning position stated above, would seem to indicate that an interactionist position, as advocated by Piaget and Inhelder (1956), would be a more appropriate position in accounting for the results of the data in this study.

It appears that children go through a sequence of form discrimination

stages. Total mean scores for the horizontal-vertical rotations were greater for both age groups than were the oblique rotation scores. This suggests that children's ability to discriminate horizontal-vertical rotations develops before the ability to discriminate oblique rotations.

It should be noted that the results of this study can be generalized only to the schools used in this study. Further research would be useful to confirm these results in a larger population. If these results are consistent with future findings, early childhood education curricula designed to teach prereading skills by using form discrimination tasks should be re-evaluated. Given that males and females may initially respond to form in neurologically different ways, it might be beneficial to teach boys prereading skills or discrimination tasks by using three-dimensional letters rather than by using the standard two-dimensional system of letters. In addition, evidence presented in this investigation tends to support an interaction theory of early childhood learning. If this theory is correct, children can most readily acquire new learning when teacher knowledge of their developmental, cognitive stage, and sex is coupled with information as to how appropriate readiness training material can best be presented.

#### References

- Bishop, C. Transfer effects of word and letter training in reading. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 215-221.
- Davidson, H. P. A study of reversals in young children. *Journal of Genetic Psychology*, 1934, 45, 452-465.
- Davidson, H. P. A study of the confusing letters, b, d, p, q. *Journal of Genetic Psychology*, 1935, 47, 458-465.
- Dick, A. O. Spatial abilities. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolinguistics*. New York: Academic Press, 1976.
- Ghent, L. Form and its orientation. *American Journal of Psychology*, 1961, 74, 177-190.
- Gibson, E. J. Learning to read. *Science*, 1965, 148, 1066-1072.
- Gibson, E. J., Gibson, J. J., Pick, A., & Osser, H. A developmental study of the discrimination of letter-like forms. *Journal of Comparative and Physiological Psychology*, 1962, 55, 897-906.
- Jeffrey, W. E. Variables in early discrimination learning. I. Motor responses in the training of a left-right discrimination. *Child Development*, 1958, 29, 269-275.
- Kolers, P., & Perkins, D. Orientation of letters and errors in their recognition. *Perception and Psychophysics*, 1969, 5, 265-269. (a)
- Kolers, P., & Perkins, D. Orientation of letters and their speed of recognition. *Perception and Psychophysics*, 1969, 5, 275-280. (b)
- Olson, D. R. *Cognitive development: The child's acquisition of diagonality*. New York: Academic Press, 1970.
- Piaget, J., & Inhelder, B. *The child's conception of space*. New York: W. W. Norton, 1956.
- Rosenberg, Z., & Birch, H. Age effects of irrelevant stimulus changes on a discrimination task by children. *Perceptual and Motor Skills*, 1969, 28, 467-475.
- Rudel, R., & Teuber, H. Discrimination of direction of line in children. *Journal of Comparative and Physiological Psychology*, 1963, 56, 892-898.
- Silver, J., & Rollins, H. The effects of visual and verbal feature emphasis on form discrimination in preschool children. *Journal of Experimental Child Psychology*, 1973, 16, 205-216.
- Welch, L. The development of discrimination of form and area. *Journal of Psychology*, 1939, 7, 37-54.
- Williams, J. P. Training kindergarten children to discriminate letter-like forms. *American Educational Research Journal*, 1969, 6, 501-514.



H . L . J A N Z E N

*The University of Alberta*

and

H . J . H A L L W O R T H

*The University of Calgary*

## Linguistic Predictors of Properties of Set

*The present study was an attempt to clarify the relational and predictable aspects of language behavior to Uznadze's (1966) concept of objectification. One hundred and eighty-six subjects were given the Uznadze set tasks and were asked to write an essay of a biographical nature. Essays were keypunched and scored by computer using Page's (1966) SCORETEXT program. Stepwise multiple-regression analysis identified that linguistic codes which embody content that is analytical and abstract are indicative of individuals who excite quickly in the Uznadze set tasks. Specifically, for those individuals who excite in the visual mode rather than in the haptic mode, the conscious act of cognitive activity suggests greater complexity, more abstraction and greater qualification in linguistic expression. The study suggests that developing greater facility in linguistic skills may well produce greater cognitive differentiation and enhance the process of objectification. (Dr. Janzen is Associate Professor in the Department of Educational Psychology, The University of Alberta; Dr. Hallworth is Professor of Educational Psychology and Director of the Computer Assisted Instructional Unit, The University of Calgary.)*

For western psychologists, the concept of set has been explained and analyzed on several dimensions. For example, Woodworth (1937) considered a *motive* to be similar to set. Dashiell (1949) refers to set as attending to particular stimuli which facilitate a response. Allport (1955) extended this behavioral definition to perception and to practically all types of behavior. According to Allport, set determines the course of learning as part of the learning process, a mechanism of drive or habit. Helson's (1964) adaptation level has been referred to as set. For him, set is a response (sometimes subjective, sometimes unconscious), an adaptation of the individual to incoming stimuli.

Uznadze (1958) conceived set, not as a mental or biological response nor, as in adaptation-level theory, a reference point for incoming stimuli; set, for Uznadze, was not a peripheral sensation, but a central, orienting, and integrating mechanism controlling all behaviour. Uznadze saw set not as a group phenomenon, but as an individual characteristic, emerging out of a need, and giving direction and purpose to every human act. Sets evolve out of a dynamic interaction of the person, his needs, and his environment. Set is a "unitary," "integrating," "orienting" mechanism, rather than a motor or mental phenomenon.

In Uznadze's theory, the primary means of halting or immobilizing set is through the process of objectification (or objectivization, as in Uznadze, 1966). As an explanatory concept, this process called "objectification" is perhaps the most central issue in Uznadze's theory. It is at the level of objectification that theoretical cognitive needs emerge. The process of objectification brings into consciousness the situation at hand. It approximates the individual's subjective representation to objective reality. As Uznadze (1961) says, "Every separate act of thought arises out of this process" (p. 161). Since thinking is usually done in a language, language becomes the means whereby man can halt a set through the process of linguistically objectifying the problem-situation, and then call in another set to action. In this way, set, language, and the process of objectification are inextricably tied. In addition, both set and language reflect (1) a person's needs and (2) the influence of the environment surrounding the individual (Bernstein, 1962a, 1966; Uznadze, 1966). The process of objectification, then, utilizing language as its means of cognitive expression, should be different for each individual, depending on the interaction of the two factors mentioned above.

It is precisely this relationship that Hertzog (1968) investigated. He postulated that the "signals" or situations to which an individual can respond will be a function of the environment found in that person's social structure. It is these signals that the individual will objectify in a cognitive relation (the process of objectification) and hence form the basis for the development of his thought and language. On this basis, the process of objectification will be different for each person, resulting in different approximations of reality and reflected in the language usage. This same concept is quite clearly related to Nunnally and Flaughner's (1963) idea of the psychological implications of individual differences in word usage. Hertzog felt that the more dynamic the psychological state of readiness for activity, the more effective will be the interaction of the individual with his environment and the more adequately will information be processed by the individual (Hertzog, 1968).

According to set theory, Hertzog predicted that individuals who excited and extinguished quickly either in the haptic or visual modality would exhibit elements of Bernstein's (1966) elaborated code. Those who excited and extinguished slowly would exhibit elements of Bernstein's restricted code. In general, this basic hypothesis was confirmed. Significant predictors of rapid excitation and extinction in the haptic and visual modalities were: (1) Adjectives, (2) Personal pronouns, (3) Abstraction words, (4) Uncommon adverbs, (5) Essay length, and (6) Verbs. High rate excitation groups, haptically or visually, distinguished themselves in the greater use of



adjectives. High rate of extinction in haptic modality groups preferred greater use of adverbs. The fast extinction group in the visual modality used greater elements of subordination. Uznadze's set theory and Bernstein's codes seemed to have some sense of theoretical verification. Reality, objectified through sets, did bear resemblance through linguistic expression, based upon cognitive differentiation. Written language did seem to reflect either a greater or lesser degree of objectification within each individual (Hertzog, 1968).

In attempting to elaborate and verify Hertzog's findings, as well as to clarify the linguistic aspects of the Soviet concept of "objectivization," the present authors felt it necessary to analyze a greater number of written essays. Hertzog analyzed eighty documents only. As can be expected, the data that must be examined in language research is so overwhelming both in quantity and complexity, that computers come as a relieving aid. In order to do this, however, certain basic assumptions have to be made. The following is a rationale for using content analysis procedures and an empirical analysis of the language categories chosen in the present study.

The technique used to analyze written documents, first developed by Stone, Dunphy, Smith and Ogilvie (1966), was content analysis. Stone argued that through the analysis of raw data, the behavioral scientist can often learn much about the personalities and preoccupations of writers. A basic assumption made in this study is that words and symbols used in written communication are indicative of many latent attributes of the source of that communication. Considerable research has established the empirical evidence that frequency counts on certain linguistic categories provide consistent relationships for personality and cognitive attributes of individuals (Newbigging, 1961; Nunnally & Flaugher, 1963; Gottschalk & Hambidge, 1955). It is only reasonable to suggest, however, that computer analysis can only rely on syntactic and lexical cues and, therefore, many denotative and connotative meanings in writing are often missed.

### *Method*

#### *Linguistic Measures*

The linguistic measures developed or borrowed for this research are shown in Table 1.

The first six categories were established by computer analysis of a total of 400 essays. A print-out of all different words (type-tokens) and their frequency were given. These words were then classified and compared with the Thorndike-Lorge (1944) word frequency list. Only those words having a frequency count of 20,000 or more were included in the language dictionaries of verbs, adjectives and adverbs. In this way, actual sample data was used to define the language category, and a more accurate and consistent word-count by computer was assured. Measures 4, 5, and 6 included all relevant words fitting into that category. The variable "Length of Essay in Words" was simply a wordcount. The remaining categories are reported in Table 1 and were used because empirical evidence existed for their validity and reliability. Measures 8, 9, and 10, for example, had reported reliability coefficients (KR-20) of .79 to .83 (Nunnally & Flaugher, 1963).

TABLE 1  
LINGUISTIC CATEGORIES DEVELOPED EMPIRICALLY

Linguistic Measures	Research Source
1. Number of Verbs (passive and active)	Measures 1 to 7 were developed empirically at the University of Calgary and published by Janzen (1973).
2. Number of Adjectives	
3. Number of Adverbs	
4. Prepositions	
5. Subordinate Conjunctions	
6. Personal Pronouns	
7. Length of Essay in Words	
8. Positive Evaluators	Measures 8, 9, and 10 were taken from Nunnally and Flaughner (1963).
9. Negative Evaluators	
10. Denotation and Abstraction Words	
11. Opinionation Words	Measures 11, 12, and 13 were taken from Hiller, Marcotte and Martin (1969).
12. Vagueness Words	
13. Specificity Words	

### *Sample*

One hundred and eighty-six subjects participated in the study. All subjects were first-year students at the University of Calgary.

Of the 186, one hundred and seventeen were females, sixty-nine were males. Since Hritzuk (1968) found no sex difference in the set variables, males and females were combined in the statistical analysis.

### *Procedure*

The independent variables were the linguistic measures determined by a written response to a suggested topic. The topic was a general one allowing maximum personal expression and freedom. No time limit was imposed. Approximate average length of an essay was 750 words.

The essays were key-punched on IBM cards and analyzed by a 360/50 computer using Page's SCORETEXT program (Page, 1966, 1967). The categories, in the order they appeared in the analysis, formed the predictor measured and were as follows:

1. Number of positive evaluation words
2. Number of adjectives
3. Number of verbs
4. Number of adverbs
5. Number of personal pronouns
6. Number of subordinate conjunctions
7. Number of prepositions
8. Number of negative evaluation words
9. Number of abstraction words
10. Number of opinionation words
11. Number of vagueness words
12. Number of specificity words
13. Length of essay in words.



All measures were standardized by dividing the count by essay-length and multiplying by a constant.

The dependent variables were the characteristics of set. To test for set properties the following equipment was used:

- (a) tachistoscope.
- (b) two slides: one slide had two unequal circles, 30 mm. and 15 mm. in diameter; the other slide had two equal circles of 22.5 mm. in diameter.
- (c) Three wooden spheres with handles. One sphere was 100 mm. in diameter, the other two spheres were 70 mm. in diameter; each sphere had a total weight of 300 grams.

The gathering of the set data involved individual testing. To test for a set in the haptic modality, the subject was seated with hands resting on the thighs and palms upward. He was blindfolded and the following instructions were then given:

I am going to present two spheres to you, one to each hand. You may grasp them for a moment and then I shall remove them. I shall repeat this a number of times. Each time I do this, please tell me if the spheres feel equal or unequal in size. If you think that they are unequal in size, tell me in which hand you feel the larger one. Please be sure to tell me each time I do this.

Following this, the two spheres of unequal size were placed in the palms, with the larger sphere placed in the right palm (setting tests). The subject was allowed to grasp the spheres and then they were removed. Then the test for set was made. This was done by presenting the two equal spheres without informing the subject (critical test). If they appeared unequal, a set had been fixated. If no fixation appeared (i.e., they felt equal in size), then the setting trials were resumed. Critical trials followed after every setting trial until a set had been fixated, or to a maximum of 20 setting trials. The verbal response given after each trial was recorded.

When the subject perceived the two equal spheres as unequal, a set had been established. To extinguish, the critical trials (presentation of equal spheres) were continued until the spheres were perceived periodically for five consecutive trials or to a maximum of 20 critical trials. The response after each critical trial was recorded.

To test for set in the visual modality, the subject was asked to focus her eyes on the binocular part of the tachistoscope. The instructions were:

You will be shown two circles briefly. Each time you see the circles I want you to tell me if the circles appear to be equal or unequal in size. If you think they are unequal, tell me which circle appears larger to you. Please be sure to tell me each time I flash circles on the screen.

The slides with the circles on them were presented for 0.1 second. The setting trials now followed the exact procedure as described above. The response after each trial was recorded.

When the subject perceived the equal circles as unequal, a set had been fixated. To extinguish, the critical trials (presentation of equal circles) were continued until the circles were perceived veridically for five consecutive trials, or to a maximum of 20 critical trials. The response after each trial was recorded.

To test for intermodal transfer, the subject was asked to look continuous-

ly into the tachistoscope and at the same time to place his hands on the table in front of him with the palms upward. The instructions were:

I am going to place the spheres in your palms again, just as I did a few minutes ago when I blindfolded you. I am also going to flash the circles on the screen every so often. However, this time you do not have to tell me what the spheres feel like. I want you to keep your eyes on the screen and whenever you see the circles, tell me if they appear equal or unequal in size. If they appear unequal to you, tell me on which side the larger one appears. Please be sure to tell me each time I flash the circles.

One setting trial with the unequal spheres was then given with the larger sphere placed in the right hand. This was followed by the presentation of the equal circles for 0.1 seconds by means of the tachistoscope (critical test). If the two circles appeared unequal, then the set fixated in the haptic sphere had been transferred to the visual modality. If the spheres were seen as equal, the setting trials were continued in the haptic sphere and the critical tests in the visual modality subsequently after every setting trial until a set had been transferred or to a maximum of 20 trials. The responses were recorded.

The procedure described above is similar to that used by Uznadze (1958). The cut-off scores used in the set tests were derived from Hertzog (1968) and Hritzuk (1968).

Each subject was scored according to the number of trials required to fixate and extinguish a set in the haptic and visual modalities, and the number of trials before irradiation occurred. The following set categories were obtained:

- A. Rate of excitation in the haptic modality
  - (1) Fast (1-5 trials)
  - (2) Slow (6-21 trials)
  - (3) None (no excitation)
- B. Rate of extinction in the haptic modality
  - (4) Fast (same criterion as above)
  - (5) Slow
  - (6) None
- C. Rate of excitation in the visual modality
  - (7) Fast
  - (8) Slow
  - (9) None
- D. Rate of excitation in the visual modality
  - (10) Fast
  - (11) Slow
  - (12) None
- E. Number of trials before irradiation took place from the haptic to the visual modality
  - (13) Fast
  - (14) Slow
  - (15) None.

It is important to note that these measures were dichotomized to allow a discrete analysis of the speed of excitation, extinction, and irradiation, with significant linguistic variables. The categorization procedure was similar to that used by Hritzuk (1968).



Data Analysis

For the data analysis, a step-wise multiple-regression program operated by the University of Alberta Research Division was used. No prior entry and deletion levels were set for criterion and predictor variables. Means, standard deviations and Pearson *r* correlations were computed.

Results

In the regression analysis all linguistic measures were entered for each of the fifteen criterion set characteristics. The analysis produced five set measures that showed significant relationship to the language variables.

Table 2 reports the step-down P values for all statistically significant (.05 level) predictor measures for the criterion "Fast Excitation in the Haptic Modality". For the thirteen linguistic measures, all but five showed a significant predictive relationship to the criterion. The variables in this regression equation indicate that the significant predictors are not function words, but content words. The group of language measures produces no clear-cut picture of the process of objectification. Suffice it to say that those individuals who excite *quickly* in the haptic modality exhibit greater action in writing (verbs), tending to self-reference (pronouns) and abstraction, whether specific or not.

TABLE 2  
STEP-WISE REGRESSION WITH ALL LINGUISTIC VARIABLES PREDICTING  
THE CRITERION "FAST EXCITATION IN THE HAPTIC MODALITY"

Linguistic Measures	Step-Down F	Step-Down P
Negative Evaluation Words	5.01	.026
Vagueness Words	3.10	.047
Pronouns	2.68	.048
Opinionation Words	2.86	.024
Adverbs	2.65	.024
Specificity Words	2.36	.032
Verbs	2.13	.041
Denotation-Abstraction Words	1.95	.054

All predictor variables were entered. Restriction levels for entry and deletion in the regression equation were not set.

The other set variables that showed significance in the analysis were: (1) Slow Excitation in the Haptic Modality and (2) No Excitation in the Haptic Modality.

For these two criterion measures, four language categories proved significant predictors. From Table 2 it is seen that no new variables were entered in this equation that did not appear for the previous one. For the criterion "Slow Excitation," however, the three language measures seem to indicate that those individuals who excite slowly, or not at all for that matter, write predominantly in a qualifying and vague way (verbs, adverbs, vagueness). These variables might indicate that, in Uznadze's terms, people who excite slowly or not at all do not respond in a dynamic manner (Unadze, 1966). Their cognitive style is static, less personal, and, in general, possesses a more limited approximation of reality (Hertzog, 1968).

TABLE 3  
STEP-WISE REGRESSION WITH ALL LINGUISTIC VARIABLES PREDICTING  
EXCITATION IN THE HAPTIC MODALITY

Linguistic Measures	Step-Down F	Step-Down P
<u>Slow Excitation</u>		
Adverbs	6.54	.011
Verbs	3.74	.025
Vagueness Words	2.73	.044
<u>No Excitation</u>		
Opinionation Words	3.74	.054

TABLE 4  
STEP-WISE REGRESSION WITH ALL LINGUISTIC VARIABLES PREDICTING  
THE CRITERION "SLOW EXCITATION IN THE VISUAL MODALITY"

Linguistic Measures	Step-Down F	Step-Down P
Adjectives	5.10	.024
Positive Evaluation Words	3.89	.022
Prepositions	3.71	.012
Subordinate Conjunctions	3.40	.010
Opinionation Words	2.94	.013
Pronouns	2.61	.018
Negative Evaluation Words	2.32	.026
Vagueness Words	2.09	.038
Adverbs	1.91	.052

Table 4 reports the only visual set category that had significant linguistic predictors.

Slow excitation in the visual modality has almost similar predictable attributes as excitation in the haptic mode. The differentiating variables, as evidenced in Table 4, are those of adjectives, prepositions, and subordinate conjunctions. This indicates that, as in the haptic mode, those who excite quickly in the visual mode have very diverse linguistic cognitive expression, but a greater tendency to complexity in writing. The complexity factor can be identified by the unique additional predictor variables. Evidence by Bernstein (1966), Lawton (1963, 1964), Page (1967), Percival (1966) and others indicates that those linguistic elements (preposition, subordinate conjunctions) are markers or proxies of elements considered "complexity in writing." Furthermore, the inclusion of these function words, by nature being resistant to change, might indicate constructive thinking, and objectification process tending toward contemplation.

The final significant set criterion was "Slow Irradiation from the Haptic to the Visual Modality." Results are reported in Table 5. Again, the usual eight language measures are entered. There is one unique element here, however, and that is the entry of the variable "length of essay in words." Uznadze (1966) states that the evidence of irradiation indicates the "transmodal nature of set," a single set expressing itself in different modalities. The variable "essay length" has been said to indicate a richness



TABLE 5  
STEP-WISE REGRESSION WITH ALL LINGUISTIC VARIABLES PREDICTING  
THE CRITERION "SLOW IRRADIATION FROM THE HAPTIC TO THE  
VISUAL MODALITY"

Linguistic Measures	Step-Down F	Step-Down P
Negative Evaluation Words	5.02	.026
Specificity Words	4.00	.019
Pronoun	3.48	.016
Adverbs	3.07	.017
Prepositions	2.58	.027
Length of Essay	2.33	.034
Subordinate Conjunctions	2.14	.041
Denotation-Abstraction Words	1.96	.053

of production, a richness of the verbal planning and translation act, a sign of an individual possessing the “elaborated code” (Bernstein, 1962a, 1966; Page, 1967; Hertzog, 1968; Janzen, 1973). Apart from the other linguistic attributes in visual and haptic excitation, the fact that essay length appeared here supports Uznadze’s contention that individuals who show irradiation express themselves in a complex and elaborate way. Perhaps this leads one to believe that such individuals possess a more complex conceptual hierarchy (Ausubel, 1965; Hertzog, 1968).

One further finding warrants mentioning. The fact that “Number of adverbs” was the only predictor common to excitation, haptically and visually, and to the irradiation measure indicates support of Hertzog’s (1968) finding that this same variable demonstrates the existence of an elaborated code. Use of adverbs has also been found to demonstrate a discriminating and intensifying quality of writing (Lawton, 1964; Cliff, 1959).

*Discussion and Conclusions*

In the Uznadze theory of sets, the concept of objectification is referred to as an act that leads to the solution of a specific problem and a new set toward that problem. Since the activity of objectifying is linked to language (Uznadze, 1966; Hertzog, 1968), it was assumed that specific aspects of style or vocabulary usage in writing would elaborate on the relationship between activated attributes of sets and their linguistic correlates. Uznadze presents no empirical evidence relating the act of objectification to such specific language factors.

For Uznadze, this internal preconscious state known as *set* exists in the subject as an integral state. It is not, according to Uznadze, an isolated mental quality to be contrasted with other conscious qualities. It is the determining factor in the dynamics of the mind. The illusions are manifestations of the activity of the set. Because sets arise out of the individual’s needs and his environment, one can safely assume that mediation is involved for the occurrence or altering of a set (Uznadze, 1966). Mediation through the process of objectification, implying language thereby, becomes one way of clarifying how individuals accommodate previous experiences, and how it is evidenced through one specific medium, namely writing.

The present study has shown that the excitation and irradiation of illusions in haptic and visual modalities have some unique linguistic attributes. Those individuals who excite quickly in the haptic modality have a tendency toward action (verbs being a unique predictor for this group of individuals). This may mean that those individuals have a process of objectification that is alert and active. Their reality is objectified through language that depicts excitement, activity, and perhaps emotional stability.

On the other hand, for those individuals who excite slowly in the visual sphere, the predominant linguistic attribute of objectification here is the use of adjectives. According to Katter (1964) and Lawton (1963, 1964), this would indicate a cognitive differentiation predisposed to analysis and evaluation. Also, since adjectives signify personal reactions to objects, it indicates a quality of feeling and emotional explicitness. The way these individuals would activate and change a set through the process of objectification would be quite different from that of individuals who excite quickly haptically. For the visual mode, the conscious act of cognitive activity suggests greater complexity, more abstraction, more qualification before new sets for new activity are aroused.

The unique linguistic attribute of irradiation has already been referred to. The variable "length of essay in words" was the defining characteristic. In Uznadze's research, many subjects would not irradiate a set from one modality to another modality. Since Uznadze found the presence of irradiation from preschoolers to adults, obviously various degrees of complexity or internalization of the word by an individual are possible. The fact that "length of essay" is the unique linguistic attribute of this phenomenon points to the availability within the subject of a greater range of linguistic symbols from which to draw. It could also imply that the more sensitive a person's nervous system is, the greater the degree of irradiation, the greater the internal store of linguistic signals.

Developing a wide capacity in linguistic skills may well serve to produce greater cognitive differentiation and enhance the process of objectification (Ausubel, 1965; Witkin, Goodenough & Karp, 1967). Furthermore, since sets toward an activity can be helpful in producing problem-solving behavior (Uznadze, 1966) and since language and thinking are intimately related (Carroll, 1964), perhaps the development of linguistic skills would further an individual's ability to process information. Through the medium of language instruction, therefore, the development of the process of objectification would be enhanced, allowing the student to stop, think, and re-experience reality.

#### References

- Allport, F. H. *Theories of perception and the concept of structure*. New York: Wiley, 1955.
- Ausubel, D. P. Cognitive structures and the facilitation of meaningful verbal learning. In R. C. Anderson & D. P. Ausubel (Eds.), *Readings in the psychology of cognition*. New York: Holt, Rinehart & Winston, 1965, 103-115.
- Bernstein, B. Linguistic codes, hesitation, phenomena and intelligence. *Language and Speech*, 1962, 5, 31-46.(a)
- Bernstein, B. Social class, linguistic codes and grammatical elements. *Language and Speech*, 1962, 5, 221-234.(b)



- Bernstein, B. Elaborated and restricted codes: Their social origins and some consequences. In A. G. Smith (Ed.), *Communication and culture*, New York: Holt, Rinehart & Winston, 1966, 427-441.
- Carroll, J. B. *Language and thought*. Englewood Cliffs: Prentice-Hall, 1964.
- Cliff, N. Adverbs as multipliers. *Psychological Review*, 1959, 66, 27-44.
- Dashiell, J. F. *Fundamentals of general psychology*. Boston: Houghton Mifflin, 1949.
- Gottschalk, L. A. & Hambidge, G. Verbal behavior analogies: A systematic approach to the problem of quantifying psychologic processes. *Journal of Projective Techniques*, 1955, 19, 387-409.
- Nelson, H. *Adaptation-level theory*. New York: Harper & Row, 1964.
- Hertzog, R. L. Set characteristics of linguistic order. In A. Prangishvili (Ed.), *Eksperimental'nye Isseldovania Pol Psikhologic Ustanovski*. Tbilisi: Academy of Sciences, 1968.
- Hiller, J. H., Marcotte, D. R., & Martin, T. Opinionization, vagueness and specificity dimensions: Essay traits measured by computer. *American Educational Research Journal*, 1969, 2, 271-284.
- Hritzuk, J. A comparative and experimental application of the psychology of set. In A. Prangishvili (Ed.), *Eksperimental'nye Isseldovania Pol Psikhologii Ustanovki*. Tbilisi: Academy of Sciences, 1968.
- Hritzuk, J. & Janzen, H. L. A Soviet and Western concept of set. *Alberta Journal of Educational Research*, 1970, 16, 217-225.
- Janzen, H. L. An investigation of the integral nature of the Soviet concept of set. *International Journal of Psychology*, 1973, 7(4), 207-218.
- Katter, R. V. Language structure and interpersonal commonality. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 132-138.
- Lawton, D. Social class differences in language development: A study of some samples of written work. *Language and Speech*, 1963, 6, 120-143.
- Lawton, D. Social class language differences in individual interviews. Unpublished manuscript, University of London Institute of Education, 1964.
- Newbigging, P. L. The perceptual reintegration of frequent and infrequent words. *Canadian Journal of Psychology*, 1961, 15, 123-132.
- Nunnally, J. C., & Flaughner, R. L. Psychological implications of word usage. *Science*, 1963, 140, 777-783.
- Nunnally, J. C., & Flaughner, R. L. Correlates of semantic habits. *Journal of Personality*, 1963, 31, 192-202.
- Nunnally, J. C., Flaughner, R. L., & Hodges, W. F. Measurement of semantic habits. *Educational and Psychological Measurement*, 1963, 17, 419-434.
- Page, E. B. Grading essays by computer, *Phi Delta Kappan*, January, 1966, 238-243.
- Page, E. B. Statistical and linguistic strategies in the computer grading of essays. Paper read at the International Conference on Computational Linguistics, Grenoble, France, August, 1967.
- Percival, E. The dimensions of ability in English composition. *Educational Review*, 1966, 18, 205-212.
- Stone, B. J., Dunphy, D. C., Smith, M. S., & Ogilvie, D. M. (Eds.). *The general inquirer: A computer approach to content analysis*. Cambridge, Massachusetts: M.I.T. Press, 1966.
- Uznadze, D. N. *Experimental investigations in the psychology of set*. Tbilisi: Academy of Sciences, 1958.
- Uznadze, D. N. *Experimental basis of the psychology of set*. Tbilisi: Academy of Sciences, 1961.
- Uznadze, D. N. *The psychology of set*. Consultants Bureau, New York, 1966.
- Witkin, H. A., Goodenough, D. R., & Karp, S. A. Stability of cognitive style from childhood to young adulthood. *Journal of Personality and Social Psychology*, 1967, 7, 291-300.
- Woodworth, R. S. Situation and goal set. *American Journal of Psychology*, 1937, 50, 130-140.

L . D . NELSON

and

T . E . KIEREN

*The University of Alberta*

## Children's Behavior in Solving Spatial Problems

*This is a report of the observed behavior of children in the age range three to eight years while solving a series of spatial problems. Elements of a cube, a regular tetrahedron, and a regular dodecahedron were constructed of Plexiglas and Velcro so that the elements could be easily assembled to their three-dimensional form and disassembled to a two-dimensional form. Children in this range displayed a variety of behaviors while attempting to form the three-dimensional shapes, but most of them were successful. There was a distinct reluctance on the part of about half the children at each age range to make predictions about the foldability of given two-dimensional layouts. Children are good builders but do not want to say in advance what form their structures will take. (Dr. Nelson is Professor in the Department of Elementary Education and Dr. Kieren is Professor in the Department of Secondary Education, The University of Alberta.)*

Developing the ability to solve problems has always been a primary aim in mathematics instruction at any level. In today's complex and changing world, that objective is probably more important than it has ever been. Since problem-solving activities such as grouping, classifying, constructing, and exploring are characteristic of young children, it would appear that problem-solving can be an appropriate aim, as well as a useful vehicle, for learning mathematics in early childhood.

What, precisely, should be the problem-solving goals for young children? Clearly, the focus cannot be on the solution of story problems. These almost invariably, unless they are trivial, involve a considerable sophistication in the use of the symbols of mathematics which young children do not have.



The goal must be oriented to the solution of real problems which can be given a concrete existence for the children.

But how can such problems be constructed? Is there a catalogue of behaviors which characterize children's activities when they face such problems? Are the behaviors age-related? Does such problem-solving experience transfer to other situations and does it affect how the child thinks about new situations? Nelson and Sawada (1975) have attempted to find the answer to these and related questions in a series of cross-sectional and longitudinal studies.

This report represents an analysis of the cross-sectional data of one problem task included in these studies. Other studies in the series focus on problems of sequences, operations, coordinates, factors, and reflections in a plane. It is hoped that analysis of children's behavior in these situations will provide further insight into the highly complex human behavior called mathematical problem-solving and will offer suggestions for the construction of problem-solving curriculum items for young children.

### *Related Research*

Despite a long history of interest in problem-solving and the current emphasis given to it in mathematics instruction, problem-solving behaviors of young children—or any other age group for that matter—have not been carefully studied (Kilpatrick, 1969). There has, of course, been the remarkable work of Polya (1957) in which he tried to analyze problems and problem sequences in terms of the thinking processes which he conjectured were needed to solve them. Although based upon years of observation and insightful teaching, Polya's work suggests how the problem-solving process *should* proceed rather than how children actually proceed when working on a problem.

Piaget and Inhelder (1967) discussed at length the attempts of subjects aged five to sixteen to represent their concept of what shape various solids would assume "if we open it out flat on the table" (p. 274). While clearly related to the work reported here, the purpose and hence the focus of this Geneva research was on the representation of space. While the current study makes use of similar geometric objects, the child was engaged in predicting the constructibility of solids and in actually constructing them. Hence, the motor skills needed for representation were not required.

Genkins (1975) reports a study of subjects in kindergarten and grade 2 as they learned the concept of bilateral symmetry under two instructional conditions. She found that younger children were more successful using paper folding rather than mirror techniques and that they generally performed at a lower level than their grade 2 counterparts. While these results have some implications for the technical aspects of geometric problem-solving, her work relates only tangentially to the present investigation. The Genkins study focussed on learning a particular geometric concept, symmetry, while the study reported here is concerned with the free responses of children to geometric problems.

It was not the purpose of this study to generate a theory of problem-solving (although it did attempt to validate a concept of "good" problems advocated by Nelson and Kirkpatrick [1975]). It was not designed to test a

theory of space perception or conception, nor did it focus on geometric concept learning. It did attempt to provide data and analysis upon which to base future development of geometric problem-solving objectives for young children.

### *Method*

To help answer the kinds of questions asked at the outset, it was necessary first to provide the child with some means of representing stable three-dimensional solids which could be readily disassembled into two-dimensional networks and as readily re-assembled to their three-dimensional form. For the purposes of this investigation, the regular solids chosen for representation were the cube, the tetrahedron, and the dodecahedron.

To fulfil the requirements of easy assembly and disassembly, the elements of the solids were cut from  $\frac{1}{8}$ -inch Plexiglas. There were six square shapes for the cube, four equilateral triangular shapes for the tetrahedron and 12 regular pentagons for the dodecahedron. The sides of these shapes all measured 10 centimetres. Alternate faces of the material called Velcro were then glued to adjacent sides of each of these elements. Velcro comes in strips with two facing sides, one of which is a plushlike fabric and the other a material consisting of a mass of tiny hooks. When the two sides come in contact, the hooks engage themselves in the softer fabric to form a join of the two sides. The photograph shows the structure of some Plexiglas shapes with the attached Velcro. A series of tasks suitable for children was devised around this apparatus.





A sample of 60 children, 10 at each age from three to eight, was chosen from a list of 200 volunteer subjects, on the basis of the availability of the subjects during the vacation period. The apparatus was placed on a low table at which each child in turn was seated. An interviewer sat beside the child and presented the tasks associated with each problem according to a pre-arranged and pretested protocol. All behavior of the child as he worked through the tasks was recorded on 1/2-in. video tape. During the session, each child worked six different problems with six different problem apparatus. The construction of several three-dimensional shapes from the apparatus described and presented above represented one problem situation. The behavior of children as they constructed these shapes was analysed and is reported here. Reports on the other problems are now in preparation.

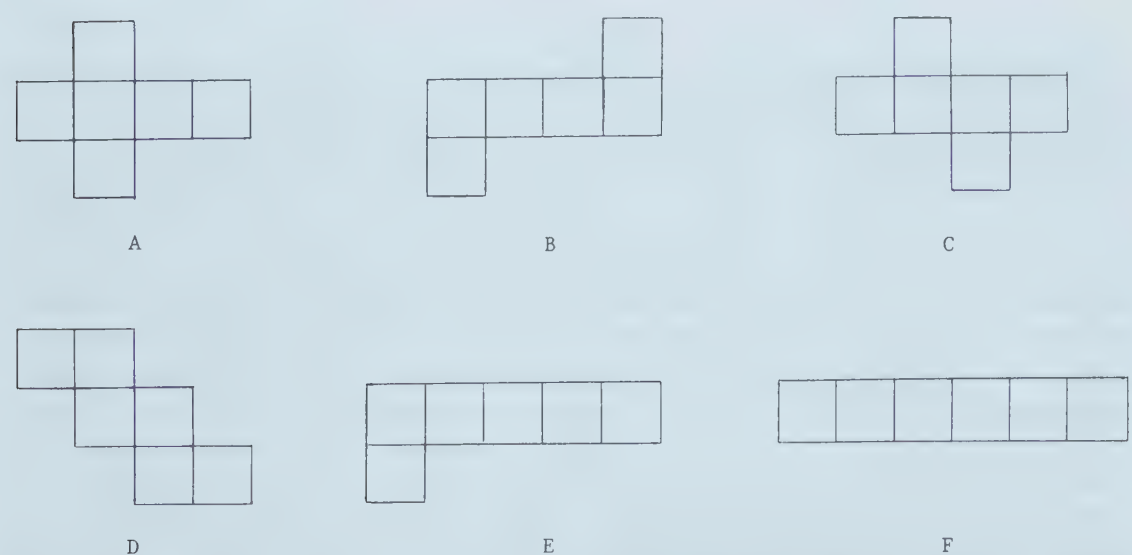
Once the data were on tape, the behaviors were transcribed in structured code to a prepared form and then carefully analysed. Sample transcriptions may be found in Little (1976).

The tasks associated with the problem apparatus described above were as follows:

*Cube*

The child was shown the assembled Plexiglas-Velcro cube and asked what he called the shape. Whatever name the child chose to associate with the shape was used in subsequent questions. He was shown how to disassemble the cube, then asked to reassemble it. A wire model of a cube was on the table for reference but no effort was made to draw the child's attention to it. As soon as the child had mastered the mechanics of assembly and disassembly, he was presented with the layouts in Figure 1 one after the other in the order shown.

FIGURE 1  
ALTERNATE CUBE LAYOUTS



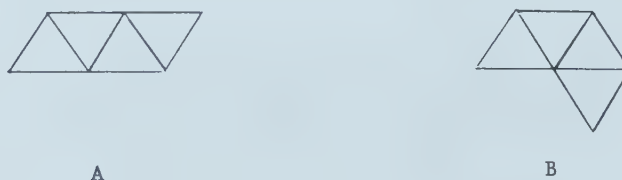
As each was presented, the child was asked if he thought the particular pattern would fold into a cube. An attempt was made by the interviewer to elicit a prediction from the child before folding began, but if the child showed any resistance, he was allowed to proceed with the task of folding.

Children who were unsuccessful in folding the first two shapes were permitted to abandon the remainder of the tasks. Layouts E and F could not be folded unless some change was made in the arrangement. The child was asked to reorganize the squares so that he could make the cube.

### *Tetrahedron*

As soon as the six tasks with the cube were completed, the child was shown the four Plexiglas-Velcro equilateral triangle shapes and asked if he could put them together to make a box. A wire network of the same shape with the same dimensions was placed on the table in front of the child. If he needed assistance in constructing a tetrahedron, it was given to him. Then the two layouts in Figure 2 were presented successively in the order shown and the child was asked if he could fold each one to make a box. A prediction was expected, but if the child resisted, he was permitted to proceed with folding.

FIGURE 2  
ALTERNATE TETRAHEDRON LAYOUTS



### *Dodecahedron*

This shape was presented to the child in assembled form and after the child had examined it, he was asked to take it apart so it would lie flat on the table. The child was then asked to reassemble the shape. There was no wire model provided in this case. If necessary, the interviewer helped to hold up the sides of the incomplete form for the child.

### *Results*

Observations of children's behavior in the shape-folding tasks will be divided into three sections: one for the cube, one for the tetrahedron, and one for the dodecahedron. The reader should keep in mind that these are observations carefully made on a small sample of 60 children. These children were in the age range three to eight years and were not randomly selected. There were the same number of children in each age range. It is not the purpose here to present exhaustive tabular data. Such data has been provided by Little (1976). This report is concerned with the qualitative nature of children's behavior rather than with reporting quantitative information.

### *Cube*

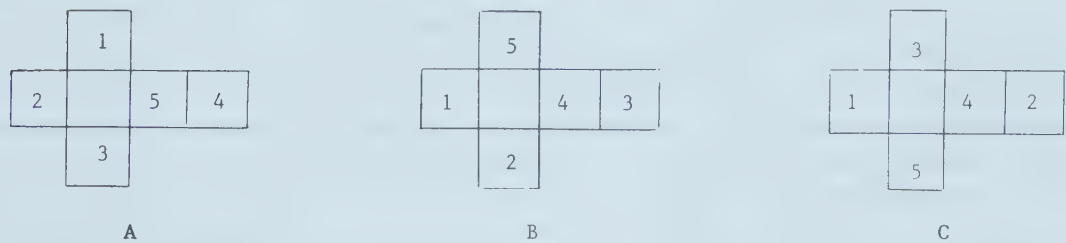
Even three-year-olds were able to construct a cube from a complicated looking two-dimensional layout. Nine out of 10 readily fit the elements of the cube corner-to-corner and edge-to-edge. Two three-year-olds and a six-year-old misinterpreted the question and piled the elements one on top of the other to make a "wallet," but 70% or more of the children correctly interpreted the tasks and successfully folded the cube. Although the older



children needed less assistance than did the younger ones, there was at least one child in each age group who needed some kind of assistance from the interviewer to complete the task.

Despite the rather intricate series of folds, turns, and fits, most children appeared to have developed some systematic procedure for handling the tasks. For instance, the most frequently used procedure for folding shape A was that shown in Figure 3 (A). The numerals indicate the order in which the elements were joined. As indicated, they would join elements 1, 2, and 3 to form kind of half box and then close it up by lifting and folding up 4, and then lifting 4 and 5. There were less systematic procedures used, of course. A few children folded in what appeared to be completely random order. Some younger children tended to work from both ends simultaneously in what appeared to be a global bringing together of the elements. These behaviors manifested themselves as indicated in Figure 3 (B) and (C). Eighteen of the three-, four-, and five-year-olds used either of these procedures while only three of the six-, seven-, and eight-year-old children employed the procedures. Only four out of the 60 children used a procedure that would start out with an unstable structure, i.e., a tunnel. As age increased, children tended to use more productive procedures.

FIGURE 3  
FREQUENTLY USED CUBE PROCEDURES



Before children were permitted to proceed with the folding, they were asked to make a prediction about whether or not they thought a layout would fold into a cube. Roughly 50% of the children refused to predict. The percentage decreased somewhat with age, but only 6 out of the 10 eight-year-olds would readily make predictions. This prediction behavior, or lack of it, applied to all the layouts. Further, this behavior did not change as children had several cube layout experiences.

Of those children who would make predictions, most were correct, with the percentage of correct responses increasing slightly with age.

There are interesting observations of the behavior of children when passed from the four foldable layouts to the first and then to then second unfoldable layouts. Only one child younger than seven was able to predict correctly that shape E, the first unfoldable, had that property. However, the unfoldability of shape F was correctly predicted by 30 to 40% of the children up to age six and by the far the majority of the older children. Indeed, all eight of the eight-year-olds who made a prediction on shape F made a correct prediction.

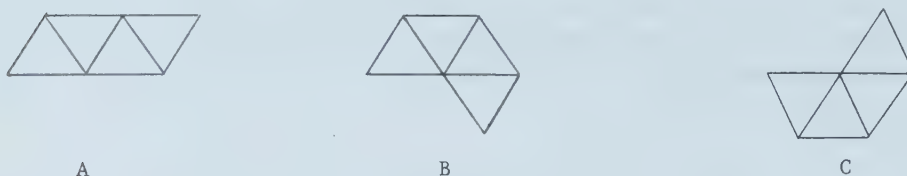
Once children found that layout E and layout F were unfoldable, they were asked to modify the layout so that it would fold. About half the

children readily modified layout E into foldable form and three-quarters were able to do so with layout F.

### *Tetrahedron*

Although to an adult the folding of four triangles into a tetrahedron may appear to be a simple task, it was not so for these children. They found it much harder to make than the cube. When they were shown four separate Velcro-Plexiglass triangles and asked to put them together to form a box, 28 of the 60 could do so. Eighty percent of the children could construct a tetrahedron from the layout shown in Figure 4 (A). However, when asked to modify the layout shown in Figure 4 (B) so that it would fold into a tetrahedron, only 20% could do so.

FIGURE 4  
FREQUENTLY USED TETRAHEDRON PROCEDURES



When asked to predict the foldability of the two layouts for the tetrahedron, there was a dramatic lack of response. Three-year-olds and eight-year-olds were the only age categories apparently willing to make predictions in the situations. About two out of three children in these categories would predict. For the layout in Figure 4 (C), none of the five-year-olds would make a prediction. Prediction responses among four-, five-, and seven-year-olds was not much better.

The nonrectilinear nature of the fit of the triangles necessary in constructing the tetrahedron seemed to be the main source of difficulty for these children. They would almost invariably align the triangles perpendicular to each other. The construction appeared often to be the result of random movements resulting finally in correct fits.

### *Dodecahedron*

The solution to this task was more age-dependent than were the solutions for the cube and tetrahedron tasks. Eight out of the 15 children who successfully dismantled and rebuilt the dodecahedron were eight-year-olds.

There were two distinct observable behaviors in children when they were asked to take the dodecahedron apart. Thirty-four of the subjects completely detached the elements of the dodecahedron. The remaining 26 created a flat layout with no element separate from the total layout.

Again, about half the children who were asked to predict refused to do so. In the dodecahedron task, however, only 34 of the children were asked to predict.

### *Conclusions*

1. The geometric problems posed for this sample, especially those involving the cube, seemed to be appropriate for children aged three to eight. The



children were able to gain appropriate *physical experience* and “solve the problem” by constructing the desired model.

2. The younger children, particularly, did not seem to gain *mathematico-logical experience* from the problems. They were less able to cope with unfoldable shapes than were older children (eight-year-olds). Although most children were good builders, the younger children tended to use less efficient folding procedures, probably indicating a lesser or different use of analysis.
3. A majority of the children in this sample, including nearly half of the older children (seven- and eight-year-olds), indicated a lack of operational control over the problems in that they refused to predict the foldability of the various layouts. This willingness and accuracy of prediction changed somewhat with age. The eight-year-olds seemed most nearly in control of difficult situations, e.g., unfoldable cube layouts and the dodecahedron problem.
4. The noncube problems presented perceptual and conceptual problems to many students. Part of this probably was due to lack of familiarity with the noncube objects. However, inability to work with nonrectilinear lines seemed a major cause of problems.

The curriculum builder, in developing geometric experiences for young children, should expect them to be motivated and willing workers. However, it should not be expected that children, especially before the age of seven or eight, will go beyond physical experience with geometric objects.

Although all the questions posed at the outset have not been answered definitively, there are some comments that should be made. In the first place, realistic problems can be created which will provide instructional experience for many children. Second, it is possible to observe and classify problem-solving behaviors of children and these behaviors seem to be age-related. This research does not, however, give definitive answers to transfer effects nor has it shed much light on the process that enables the child to “mathematize” his environment.

The purpose of the research was to gather and analyse carefully young children’s reactions to selected geometric problem-solving situations. Conclusions 1 to 4 above were not looked for in advance. Hence this research and its data, while generating and supporting these conclusions, were not explicitly designed to test these conclusions. This is the work of further research.

This research was supported by Canada Council research grants S73-0488 and S74-1206.

### *References*

- Genkins, E. The concept of bilateral symmetry in young children. In M. F. Roszkopf (Ed.), *Children’s mathematical concepts*. New York: Teachers’ College Press, 1975.
- Kilpatrick, J. Problem solving in mathematics. *Review of Educational Research*, 1969, 39, 523-534.
- Little, J. A mathematical and cognitive analysis of children’s behavior in spatial problems. Doctoral dissertation, University of Alberta, 1976.
- Nelson, D., & Kirkpatrick, J. Problem solving. In J. Payne (Ed.), *Mathematics learning in early childhood*. Reston, Va.: NCTM, 1975.

- Nelson, D., & Sawada, D. Studying problem solving behavior in young children—Some methodological considerations. *Alberta Journal of Educational Research*, 1975, 21(1), 28-38.
- Piaget, J., & Inhelder, B. *The child's conception of space*. New York: W. W. Norton & Co., 1967.
- Polya, G. *How to solve it*. New York: Doubleday & Co. Inc., 1957.



SHELLEY PHILLIPS

University of New South Wales

## The Contributions of L. S. Vygotsky to Cognitive Psychology

*An examination of Vygotsky's theory shows close affinity with western developmental-type cognitive psychologies. The compatibility of the two approaches inevitably grows out of the similarity of their basic assumptions. For example, in both approaches the Darwinian and functionalist influence is paramount, the study of cognitive processes is seen as a worthwhile scientific endeavour, and there is an anti-behaviourist stance. Vygotsky figured strongly in the demise of behaviourism and the rewriting of psychology in the Soviet Union. His ideas are now basic to Soviet cognitive psychology. Apart from discussion of these points, examination of Vygotsky's work reveals many promising concepts which could be of interest to psychologists generally, although his contributions have not been treated extensively in western psychological literature. (Dr. Phillips is at the School of Education, University of New South Wales in Kensington, Australia.)*

How often, despite the revelations of the post-Sputnikian period of reappraisal, do we hear Soviet psychology described as "behaviourist"? While there are many and varied uses of this term, typically it is used in this connection in the pragmatic sense and in the tradition of J. B. Watson. To the average Soviet psychologist, behaviourism in this style contradicts and strikes at his basic philosophy, yet this description of Soviet psychology has become a convention among western psychologists. In fact, behaviourism enjoyed its heyday in Soviet psychology some fifty years ago in the twenties; since then Soviet psychology has become progressively more cognitively oriented. It also has much in common with the developmental psychologies which dominated the psychological scene in Europe and America in the early decades of this century and which are lately regaining popularity in western countries. Pavlov was long ago reinterpreted in the marxist context to contribute to a psychology in which the study of consciousness of cognitive is central.

The role of L. S. Vygotsky in putting the study of cognition and thinking at the centre of Soviet psychology is paramount. His contributions to a psychological theory of cognition are worth examination, not only to put the record straight about Soviet psychology, but because Vygotsky's ideas and research are interesting in themselves and are not fully treated in our literature. Unfortunately, Vygotsky was a psychologist for not much more than a dozen years before he died of tuberculosis in 1934 at the age of 38, and much of what follows is but a brief sketch for a cognitive theory. The potentialities have not been fully exploited and are worthy of more attention by researchers in this area.

### *Influence of Lenin*

When Vygotsky entered the psychological field in the Soviet Union in the early years after the revolution, it was dominated by the reflexology of Bekhterev. Like Watson, Bekhterev (1933) believed the processes of consciousness or cognition can never be part of scientific investigation. Lenin, as head of the Soviet government at the time, was critical of Bekhterev's theories which he regarded as being in fundamental opposition to the philosophy of Marx and Engel which, together with his own writings, became the marxism-leninism which later formed the foundation of all aspects of Soviet science and politics, including psychology.

Briefly, marxism-leninism takes the study of consciousness or cognition very seriously and expects psychology to do likewise. Like many western psychologies, it states that psychic or cognitive processes cannot be reduced to physical, chemical, or physiological laws. As in western genetic psychologies, development in history, society, and the individual is seen as taking place through the dialectical process and qualitative change; development proceeds through a hierarchy of stages, each with its own special characteristics. Each transition to a new level of development constitutes a qualitative change which arises from an accumulation of changes in the previous stage; these changes, in their synthesis, completely change the nature of their own functions and organisations. Thus nothing can be considered singly; the study of interrelationships of the whole is essential. Since marxism-leninism stresses interrelationships in scientific study, it is opposed to the specificity which is an aspect of the pragmatic approach to science, and that style of behaviourism and neo-behaviourism which studies the learning of simple habits, maze running, pin ball throwing, or the reinforcement of single aspects of classroom behaviours. There is another important point which is often misrepresented in traditional views of marxism. According to Marx, man is a product of his environment, but *not* in a purely passive way. Marx was critical of mechanistic-materialistic theories of human development, and believed that by acting on the external world and changing it, man changed his own nature (Marx, 1935).

Many of marxism-leninism's basic concepts about the processes of development and change thus show a strong Darwinian influence, and have an affinity with the way these processes are described in some branches of biology or chemistry and the genetic style psychologies. The theories of John Dewey, James Mark Baldwin, Maria Montessori, and Jean Piaget are examples. Thus, when Vygotsky worked toward adapting Soviet psychology



to marxism-leninism, he was laying the foundations for an approach which is most familiar to a particular group of western developmental and cognitive psychologists. This approach is the antithesis of Watsonian and pragmatic style behaviourism.

### *Cognitive Psychology's Struggle*

Vygotsky's battle to give the study of cognitive processes a central place in Soviet psychology is very similar to that fought by cognitive psychologists in the west since the middle decades of this century. His solutions and outcomes are often similar. Sometimes they are different and this difference suggests some novel research approaches for western cognitive psychologists. First, there is the grappling with the question of whether the pragmatic methodological divisions and compartmentalisations typical of behaviourism are suitable for the study of cognition and development. Can learning and cognition be reduced to simple habits? Is learning and cognition an additive and chaining process of habit formation? Secondly, there is the much larger question of whether development and learning can be separated for the purposes of study. In the west, since Thorndike, we have witnessed the separation of learning theory and developmental psychology. The behaviourists and neo-behaviourists have favoured the study of simple learning tasks in the laboratory in keeping with their methodological view of science, and have downgraded or avoided the more complex developmental processes as a study.

### *An Interactionist View of Learning and Development*

Vygotsky (1962) came out in favour of an interactionist view of learning and development; his views, while compatible with marxism-leninism, have much in common with the work of contemporary western developmental psychologists. His first conclusions and demonstrations of the usefulness of this approach came from the practicalities of his own research and, in particular, from his studies of aphasia, the affection of the speech zone which disturbs processes of perception, voluntary attention, logical memory, and discriminatory and directional behaviour. Here he found it more useful to study functions and relationships between functions, rather than utilise habits as the units of study. For instance, in the early years of childhood, sense and perceptual functions are central; as these functions interlink, so sensori-motor thought or a higher level of intellectual functioning emerges. Later sensori-motor thought comes under the influence of a new function, namely language, and the child learns to classify, code, and think in a much swifter and more organised fashion than when he was dependent on his percepts and senses alone, although basically thought grew out of these processes.

With the advent of language, there is a qualitatively new stage of cognitive functioning which changes the old relation between perception and the senses. This, Vygotsky concluded, is why a lesion in the occipital parts of the cerebral cortex in an adult leads to comparatively limited disturbance of perception and does not affect vocal processes of intellectual activity. But if a similar localised lesion of the occipital parts of the brain occurs in early childhood, it leads to massive mental retardation. At this

stage, focal lesions destroy the sensory basis for the development of language and later verbal or abstract thought.

Vygotsky, being a dialectical materialist and a Darwinian developmentalist, is here stressing the qualitative difference between the thought of the child and the adult. In fact, what he is saying at the neurological level is substantiating what Piaget has to say at the level of organisation of knowledge. According to behaviourism, said Vygotsky (1963), one and the same mechanism underlies the formation of habits, whether in the adult or the child. But there are essential developmental differences between the process of an adult learning, for example, to use a typewriter, ride a bicycle, or play tennis, and the processes of a school age child learning written speech, arithmetic, and natural science. Learning to use a typewriter represents the establishment of a number of habits which in themselves do not change an adult's mental traits. Learning of this kind makes use of an already elaborated and completed course of development and, precisely because of this, contributes very little to general development. But the process of learning to talk, read, and write calls forth a very different and much more complex course of developmental processes and the advent of these processes signifies a general change in the child's mental organisations. Vygotsky hypothesised that these changes are interlaced with the development of the central nervous system, and that learning stimulates the internal processes of development. He advocated that the rise and fall of these internal lines of development be studied in conjunction with developmental changes in the course of school learning. The notion is exciting. Unfortunately, because of Vygotsky's early death, like many of his ideas this was but hypothesis although his students and associates, such as Leontiev, Galperin, Elkonin, and Luria, have extended his work in this area. Unfortunately, also, the concept of the dialectical process, while it has had explanatory usefulness in chemistry and biology, needs much more elaboration in psychological theory. In Vygotsky's description given above, it is but an intuitive idea.

### *The Zone of Potential Development*

In Vygotsky's second demonstration of the interaction of learning and development, there are some interesting differences from western approaches. Where western educational psychology has been extensively influenced by the old Herbartian tradition of readiness of the Piagetian concepts of assimilation and accommodation (development can do without learning), or the even more conservative maturational view (learning follows development), Vygotsky offered the zone of potential development. He rejected the view that the child develops through gradual and slow assimilation of changes. His observations led him to believe that the essence of cultural development consisted of a clash between the developed cultural forms of behaviours which the child encounters and the primitive forms which characterise his own behaviours. Cognitive development arises from the child's social interaction with those around him and collective activity leads him into realms which he cannot attain at the level of individual functioning. In his experiments Vygotsky found that young children can imitate, under adult guidance, a number of actions in collective activity which they cannot achieve independently. Vygotsky (1963) called this



divergence, between the level of performing tasks which are accessible under guidance with adult help and the level of performing tasks accessible to independent activity, the zone of the child's potential development.

The idea of the zone of potential development clashes drastically with the assumptions of the mental testing movement, which became traditional in the west. Intelligence tests can only determine the present and independent level of mental development, and educators tend to regard this as the limit which the child cannot transcend. Teaching is geared accordingly and to the yesterday of the child's development. But Vygotsky's theory of the zone of potential development holds that the only good teaching is that which outpaces development. The teacher should orient himself to what the child can achieve under collective activity, not simply to what he can do independently.

The zone of potential development which relates to collective activity in family and school does have certain variations in contemporary western cognitive and developmental theories. Piaget, for one, stresses social relations as producing qualitative changes in the child's development. For example, interactions and clashes with the varying points of view of his peers obliges the child to see himself from the point of view of others and is significant in the development of the concept of self, moral judgments, and the course from heteronomy to autonomy.

### *Social Interaction*

In stressing the role of collective activity in cognitive development, Vygotsky once more found the experimental units and procedures of the behaviourists inadequate. Instead of stimuli and response at the experimental level, he worked with more global categories such as "social influences" and "social interactions." For example, he examined the history of the pointing gesture (1966). At the beginning it is merely an unsuccessful grasping movement aimed at an object. When the mother interprets this as a pointing gesture, the situation essentially changes. The pointing gesture becomes a gesture for others. When the child associates the unsuccessful grasping movement with the entire objective situation, he begins to treat this movement as a pointing gesture. Here the function of the movement itself changes: from a direct movement directed towards an object, it becomes directed towards another person—a means of communication. The child is the last to realise the full purport of his own gesture, since meaning and function are created first by the objective situation and then by the people surrounding the child. Thus, relations expressed in cognitive processes were at one time real relations among people. We also become ourselves through others, and this rule applies not only to the personality as a whole, but also to the history of every individual function. The extension here of George Mead's theory of how the concept of self emerges is clear. The idea also owes much to that of Mark Baldwin and, together with these two theories, Vygotsky's description is one of the early attempts to demonstrate concretely the relationship between external and internal behaviour.

*The Role of Language and Significates  
The Mediated Regulation of Behaviour*

Vygotsky's holistic view of brain functioning and his view of the role of collective activity in cognitive processes must be held constantly in mind when discussing his views on the role of language and significates in human consciousness. While studying the effects of Parkinson's disease, in which affliction of the subcortical motor ganglions leads to disturbances of voluntary motion. Vygotsky noticed that patients suffering from this disease, although unable to step across a smooth floor, can go upstairs well. Vygotsky believed that this had to do with mediated regulation of behaviour by the undamaged cortex. When Vygotsky suggested that the patient throw cards on the floor and step across them, the patient walked without difficulty across the floor. Vygotsky interpreted this as the patient having created new additional stimuli in the external medium and performing in the mediated way the tasks which he was incapable of doing directly (Luria, 1934).

Thus, Vygotsky demonstrated that the use of signs could be used to form new functional systems to reconstruct the work of a sick brain and compensate for an existing defect. In this vein, Vygotsky did a number of experiments on mediated memorisation, in which a series of words were memorised with the aid of a series of auxiliary signals such as pictures of words. These aids were shown to extend the memory. In other experiments, Vygotsky found that if the child at the more primitive stages of thought is faced with a problem above his natural capacities, he will utilise extraneous objects, knots in string, etc., as signs to help in the solution.

Vygotsky quite honestly stated that he did not know how the process of signification actually works, other than by saying it involved "mastery of one's behaviour," which is a mediated process and which he felt sure could ultimately be explained by natural laws without resorting to spiritual factors such as "the will" in the sense that it was traditionally used in Vygotsky's student days. The concepts of mastery of behaviour and such constantly recur throughout twentieth century descriptions of thought processes. They have served, as in Vygotsky's system, to shift the emphasis from spiritually to materially based self-control, and are an important aspect of the development of twentieth century research into the psychology of thought. Unfortunately, apart from offering the possibility of a material explanation, they merely replaced one inexplicable concept for another.

The important role that signs, signals, and language play in Vygotsky's theory owes much to Pavlov. In fact, Vygotsky was the first psychologist in the Soviet Union to use Pavlov's ideas in a cognitive theory, but Vygotsky's views on the development of language have been well discussed elsewhere (Schmidt, 1973).

*Concepts*

Vygotsky's study of concepts demonstrates once again his dissatisfaction with stimulus and response as units for experimental work and his preference for social processes, functions, structure, and systems as the basis for analysis. For example, it was simple experiments by Vygotsky which looked at the early stages of the development of concepts via the process of



mediation and demonstrated the now well known observation that children “explain” the names of objects by their attributes. According to them, an animal is called “cow” because it has horns, “calf” because its horns are small, “dog” because it is small and has no horns. An exchange of names would mean an exchange of the characteristic feature of the object, so inseparable is the connection between them in the child’s mind. The structure of speech does not mirror conceptual thought, and grammar precedes logic (Vygotsky, 1962).

Vygotsky was inevitably critical of experimental studies of concept formation which either concentrated on the verbal definitions of children or required children to abstract common traits. He preferred the then new experimental approach of one of his contemporaries, Clark Hull, the American neo-behaviourist, who set up artificial concepts for subjects to learn by attaching a nonsense word to a particular combination of object attributes for which no ready concept or word exists. The advantage, as he saw it, was that the solution of the problem does not presuppose previous experience or knowledge on the part of the subject. Where Vygotsky’s approach differs from Hull’s neo-behaviourism is that, as indicated earlier, he was not interested in stimulus and response alone but in the investigation of the functional conditions of concept formation.

The stages of development of concepts as given by Vygotsky utilise, in part, work derived from Piaget and Claparède and are remarkably similar to Inhelder’s more recent descriptions. There is free use of Pavlov’s concepts of connections, analysis, synthesis, and generalisation, and the concept of quantitatively different stages is basic. According to Vygotsky, the young child takes the first step toward concept formation when he puts together a number of objects in unorganised heaps, in order to solve a problem which adults normally solve by forming a new concept. At this stage, the child tends to merge the most diverse elements into one unarticulated image on the basis of some chance impression. He mistakes his own subjective interpretations for real bonds between objects. This trial and error grouping, Claparède and, later, Piaget called syncretic thought.

During the next stage, the composition of the group is determined largely by the spatial position of the objects. Later the child groups things according to bonds actually existing between the objects, but these are based on concrete or factually present connections, not logical unity. This Vygotsky labelled as thinking in complexes. Then comes what Vygotsky called “chain complexes,” i.e., the child starts by collecting according to one trait and then his mind switches to another. The decisive attribute keeps changing. The single trait is not abstracted; it depends upon the perceptually concrete. The diffuse complex is a result of unlimited generalisation in his mind. The pseudo-complex is a transitional stage when the child appears to use the word correctly as a classification, but at first it has, in fact, all the tendencies of complexes.

The principle function of complexes is to establish bonds and relationships. It creates a basis for later generalisations. Synthesis is not as yet combined with analysis. There is an overabundance of connections and weakness in abstracting elements. Later the child can group on the basis of single attributes. The traits once abstracted, remain stable. These are

potential concepts. A concept emerges only when the abstracted traits are synthesised anew.

Thus, at a time when the behaviourists were presenting the first experimental examinations of concepts, Vygotsky offered an alternative to their associationist views. Concepts, he said, do not develop as a result of summing traits, but constantly alternate between the particular and the general. Concept formation is a process in which all elementary mental functions participate in specific combinations. The development is furthered and guided by the use of words as the means of actively centering attention, of abstracting certain traits, synthesising them, and symbolising them by a sign.

### *Spontaneous and Scientific Concepts*

Borrowing in part from Claparède and Piaget, while at the same time adapting psychological theory to marxism-leninism, Vygotsky decided that concepts were of two kinds, spontaneous and scientific. Children who can tie a knot or build a house, but who cannot describe how they did it, have not yet learned to isolate and relate facts. Their concepts are spontaneous. Once they are aware of the activity of their minds and can abstract true relationships, they operate with scientific concepts. Vygotsky's spontaneous concepts bear a close relationship to Piaget's definition of concrete thinking, an aspect of which is its limited classificatory ability and, in fact, Vygotsky (1962) stated that the cardinal difference between spontaneous and scientific concepts is the absence of system.

Spontaneous and scientific concepts interact in many ways, and it is this interaction which is essential to the growth of thought. The spontaneous concept prepares the way for the scientific. For example, historical concepts can begin to develop only when the child's everyday concept of the past is sufficiently differentiated, and when his own life and the life of those around him can be fitted into the elementary generalisation of in the past and now. His geographic and sociological concepts must grow out of the simple scheme "here and elsewhere."

Scientific concepts, in turn, supply a variety of structures for the upward development of the child's spontaneous concepts toward conscious and deliberate use. The development of a scientific concept begins with a verbal definition. It starts its life in the child's mind at the level that his spontaneous concepts reach later. A child's everyday concept, such as "brother," is saturated with experience. yet when he is asked to solve an abstract problem such as how many brothers his brother has, he becomes confused. On the other hand, though he can correctly answer questions about "slavery," "exploitation," or "civil war," these concepts are schematic and lack the rich content derived from personal experience. They are filled in gradually, in the course of further school work and reading. One might say that the development of the child's spontaneous concepts proceeds upward and the development of his scientific concepts, downward to a more elementary and concrete level (Vygotsky, 1962).

Vygotsky was particularly interested in the relationship between instruction or teaching and spontaneous and scientific concepts. He believed that school instruction induces the generalising kind of perception and thus



plays a decisive role in making the child conscious of his own mental processes. Reflective consciousness comes to the children through the medium of scientific concepts and the processes whereby they learn their hierarchical interrelationships.

In the scientific concepts that the child acquires in school, the relationship to an object is mediated from the start by some other concept. The very notion of scientific concept implies a certain position in relation to other concepts, i.e., a place within a system of concepts. Thus Vygotsky (1962) saw, as does Piaget, thought developing according to increasing development of generalisations built upon generalisations of the preceding level. The following example may illustrate the function of varying degrees of generality in the emergence of a system. A child learns the word "flower" and, shortly afterwards, the word "rose". For a long time the concept "flower," though more widely applicable than "rose," cannot be said to be more general for the child. It does not include and subordinate "rose"—the two are interchangeable and juxtaposed. When "flower" becomes generalised, the relationship of "flower" and "rose," as well as of "flower" and other subordinate concepts, also changes in the child's mind. A system is taking shape. Passing beyond earlier stages and the generalising of attributes which define various objects, the child reaches a stage where certain aspects of these objects can be abstracted and generalised into ideas of number, for example, while at the next stage, algebraic concepts may be used to represent abstractions and generalisations of certain aspects of number. Each of these stages signifies a new departure and a new and higher plane of thought.

#### *Vygotsky's Status*

To elaborate upon the application and extensions of Vygotsky's theory in contemporary Soviet psychology is an enormous subject in itself. Nor has a detailed account been given of his strivings and clashes, his fall and rise in popularity in the Soviet Union. That is also a lengthy topic, but a few brief comments may be of interest. During the twenties, when Bekhterev was popular and Yenchman, Frolov, Blonski, and others were advocating that psychology be replaced by physiology, the Ukrainian Commissariat of Education, at one stage, decided that reflexology was to serve as the basis for the establishment of scientific pedagogy and did away with all psychology for a limited time. In this period, a few articles and books appeared in which reflexology was condemned and the rapprochement between psychology and marxism-leninism was attempted, but theirs and Vygotsky's role at this time was somewhat similar to the role occupied by Piaget among the American behaviourists and neo-behaviourists from the late thirties to the fifties. It was not until the First All Union Congress for the Study of Human Behaviour in January, 1930, that measures were officially taken for introducing dialectics into all behavioural sciences, although it had been done earlier in other sciences. This conference officially established the principle of the irreducibility of psychic events to the laws of physiology. Vygotsky's work was influential in this achievement.

However, once the battle against behaviourism was won, Vygotsky's theories began to have difficulties in another direction. They were criticised because they did not always conform to the marxist-leninist line as it was

interpreted in the early years between the late thirties and the Pavlovian conference. Two years after his death, Vygotsky's book, *Language and Thought of the Child*, was officially suppressed and did not reappear until 1956 (Vygotsky, 1962). Why this happened is not clear but an official comment on Vygotsky states:

He tried to explain the human psyche as a product of human development, but in so doing he made a dualist juxtaposition of "cultural and natural" development and treated the historical development of the psyche abstractly and outside the class struggle. (Maxwell, 1962, p. 562.)

Despite this and similar criticisms, Vygotsky's contribution appears never to have been wholly ignored. Anan'ev's article in 1948, which seems to summarise the position of Soviet psychology during and after the war, stated that research into thought and speech in that period comprised one of the central places in Soviet psychology, and she included Vygotsky in her summary of such work. The present status of Vygotsky in Soviet psychology is considerable as can be seen by the frequent references of his work in contemporary studies.

### *Implications of Vygotsky's Work for Psychological Research*

For the prospective doctoral student and other researchers, Vygotsky is a virtual treasure trove of as yet unexplored topics. In particular, there are two interesting areas which could do with more extensive research in the west. First, there is the zone of potential development. Most teachers are not overly conversant with the developmental stages of their pupils, are subject rather than person oriented, and see achievement as a competitive and individual matter. Schools are still oriented to the concept of maturation, readiness, and testing the child's present intake. Classroom application of Vygotsky's ideas on the zone of potential development and collective activity, and empirical testing of the possibilities and outcomes seems long overdue, even if only as a means of examining traditional practices more objectively.

Secondly, Vygotsky combined physiological and neurological studies with the study of the development of thought. His idea that learning stimulates the internal process of neurological developments and that these two be studied in conjunction has received limited attention in the west. Surely it is clear that thought has something to do with the brain. At the same time as the structures of thought processes are being examined, is it not logical that such research be in close touch with research at the neurological and biochemical level? The distinct and pragmatic division of research in these two areas which has been characteristic of the twentieth century has obviously not been fruitful. We still do not know what thought processes are.

In conclusion, some current research which has been inspired by some of Vygotsky's tentative formulations is worth perusal. His views on the role of early social interaction in the development of language and thought are currently being examined by J. S. Bruner (1975) and his students at Oxford. Bruner has been much influenced by Soviet psychologists and his interchange with Vygotsky's associates has been considerable.

On the Soviet scene, Vygotsky-type experiments by L. F. Obuchova (1966)



on the function of auxiliary or mediating signs are of interest. She demonstrated that if a new mediational way of thinking were created by means of which the child could make concrete comparisons of an object before and after its dimensions were changed, conservation could be developed earlier than Piaget had suggested. Mediation was through a third object operating as a standard or measurement and, in this way, children were taught to measure length, volume, area, and weight of objects. They began to talk about changes in dimensions which made quantities appear different when, in reality, they were not. They were taught to single out different characteristics of objects and then to distinguish dimensions in which one object differed from the other and dimensions in which they were identical.

Another experiment by Teplenkaya (1966) which looked at the forming, in six to seven-year-olds, of the logical structures of class membership and relationships between classes and subclasses is also based on Vygotsky's work. It consisted of analysing the tasks to be learned into the necessary algorithm of action without losing the relationships involved in the task as a whole. These were the units for study and were conceived in the same spirit as Vygotsky's use of word meaning as the unit of verbal thought. After the analysis, the children were involved in performing the selected actions which dealt with all aspects of constant and variable features of classes such as, for example, the putting of boxes of different sizes within each other, the drawing or measuring of different parameters of the objects involved, and so on. Then the children were led through the subsequent stages seen as necessary for the development of mental actions as outlined by Vygotsky, namely logical and verbal expression of the activities involved, silent thinking, and consolidation and internalisation of action with appropriate practice.

The experiments demonstrated that children could be taught logical actions involved in class membership as early as six or seven years of age, when by traditional procedures, these are not normally acquired until twelve or thirteen years of age. Teplenkaya (1966) concluded that the usual level for forming such concepts reflected only the course of the process of learning in concrete historical conditions of teaching rather than the age potentialities of children and the necessary logic of the process of learning.

These are but a few of innumerable experiments which point to the rich source of research ideas offered by Vygotsky's work.

#### References

- Anan'ev, B. Achievements of Soviet psychologists. *Journal of General Psychology*, 1948, 38, 257-262.
- Bekhterev, V. M. *General principles of human reflexology*. London: Jerrolds, 1933.
- Bruner, J. S. The beginnings of intellectual skill: 1. *New Behaviour*, Oct. 2, 1975, 20-24.
- Bruner, J. S. The beginnings of intellectual skill: 2. *New Behaviour*, Oct. 9, 1975, 58-61.
- Luria, A. R., & Vygotsky, L. S. *Character and personality*, 1934, 3, 238-240.
- Marx, K. Thesis on Feuerbach. In E. Burns (Ed.), *A handbook of marxism*. New York: Random, 1935.
- Maxwell, R. (Ed.). *Information U.S.S.R.* Oxford: Pergamon Press, 1962.
- Obuchova, L. F. Experimental formation of representations of invariance in five and six-year-old children. *Proceedings of XVIIIth International Congress of Psychology*, Moscow, 1966, Symposium 24, Concept formation and inner action, 48-58.

- Schmidt, W. H. O. *Child development: The human cultural and educational context*. New York: Harper & Row, 1973.
- Teplenkaya, K. M. The formation of logical structures in 6-7-year-old children. *Proceedings of XVIIIth International Congress of Psychology*, Moscow, 1966, Symposium 24, Concept formation and inner action, 115-120.
- Vygotsky, L. S. *Thought and language*. New York: Wiley, 1962.
- Vygotsky, L. S. Development of higher mental functions. In *Psychological Research in the U.S.S.R.* Moscow: Progress Publishers, 1966.



J. C. POWELL

University of Windsor

## The Developmental Sequence of Cognition as Revealed by Wrong Answers

*Gorham's Proverbs Test (1957) was administered to 550 children from three different schools in grades 3 to 8 inclusive. The highest and the lowest scoring 16% were interviewed to determine the reasons for the wrong answers given. Wrong answers were then grouped by a two-stage cluster analysis procedure and the resulting clusters used to develop subtest scores. The correlation matrix (of right and wrong answers) was reordered into a Simplex pattern. The resulting pattern clearly supported the possibility of a developmental sequence. The clusters were then interpreted, using the transcribed reported reasons for answers as a guide. These interpretations produced a logical description of the apparent developmental sequence when placed in the same order as revealed by the Simplex arrangement. A comparison was made between this order and the hierarchical sequence found elsewhere with mature adults (Powell, 1970). The rank order correlation between the orders of the common elements in the two studies was .86, suggesting a strong relationship and supporting the empirical validity of the order obtained. (Dr. Powell is Assistant Professor in the Faculty of Education at the University of Windsor.)*

About fifteen years ago, the present author noticed a tendency for consistency in the selection of wrong answers on multiple choice tests. Specifically, a number of students tended to choose several wrong answers in common from several different items. Also, asking specific students why they chose these wrong answers often assisted in the planning of instruction.

These observations led to the question, "Are wrong answers really blind guesses as is commonly held?" A careful study using Gorham's Proverbs Test (1957) with adults revealed that this assumption is false (Powell, 1968).

Within the logic of test analysis, it is not sufficient that wrong answers be systematically selected. If answers are either right or wrong, then knowing that an answer is wrong adds no new information to the observation than an answer is not right. Such a "linear dependency" assumption is commonly held concerning wrong answers on multiple choice tests. This linear dependency question was explored (Powell & Isbister, 1974), using 387 adults and a different test, and was found to be false. With more than one alternative from which to choose, at least some wrong answers load on factors which contain no right answers; that is, some are not linearly dependent.

This latter discovery led to the question, "Do wrong answers contain achievement information?" The next study in the series (Powell, 1970) used the same test as in Powell and Isbister (1974). This study with 277 mature adults explored this question, with three conclusions:

1. Wrong answers may be better predictors of present and future performance than right answers (i.e., they contributed more of the variance, using stepwise multiple regression, to the prediction equation). Performance is defined in terms of obtained scores on course related achievement tests.
2. Wrong answers appear to be ordered in a hierarchy which parallels Bloom's *Taxonomy* (1956) and which influences the level of functioning of the questions in which they are embedded as distractors.
3. When the results of wrong and right answers are combined, the most stable (in cross-validation terms) prediction of present and future success is obtained.

Research up to this present study into the nature of behavior related to wrong answers may be summarized as follows:

1. The common assumption that a person writing a multiple choice test always either knows the answer or guesses blindly is false.
2. The related assumption that all wrong answers are linearly dependent upon the right answers is false.
3. The assumption that wrong answers contain no useful information (in a predictive sense) about achievement is false.

These conclusions imply that answer selection is much more complex than a know-guess dichotomy would seem to suggest. Do people choose answers based upon some systematic aspect of their personality? The "yes" answer to this question can be supported if a developmental sequence of wrong answer selection can be uncovered. It is to this latter issue, the exploration for a developmental sequence, that this present study turns.

## *Instrument*

## *Method*

Gorham's Proverbs Test (1957), in its multiple choice version, involves 40 proverbs for each of which four possible translations are offered. Although it also has clinical uses, its primary value with normal subjects is as a short, broad-spectrum, reading comprehension test. Norms are offered in the manual for the range of grade 5 through adult for the Abstract scale. Norms are not given for the Concrete scale. The wrong answer interpretations reported here were developed in the present study.



### *Procedure*

This test was administered to 550 children in grades 3 to 8 inclusive (approximately 90 children in each grade) from three different schools. The five highest scoring and the five lowest scoring children from each grade level of each school were interviewed to determine the reasons they had for the selection of the wrong answer. This interview followed the administration of the tests which were scored immediately so that the children could be brought to the examining room usually within minutes of the completion of the test. Two trained interviewers were employed in each school. In no case were the reasons collected more than two hours later than the test administration. This procedure was used to minimize the effects of forgetting.

A frequency count of the selection of each of the four alternatives partitioned by age was produced. The proportion of selection of each wrong alternative by age was calculated. It was observed that particular alternatives tended to have the highest proportion of selection at particular age levels. These were grouped together and further analyzed so that alternatives which were significantly correlated were placed into definitive clusters. A unique group was also generated from some alternatives which were bimodal in their highest proportions.

Thus age-level related groups of wrong alternatives were placed into homogeneous clusters. Those clusters with five or more members were treated as subtests and the answers of all 550 children were rescored. Those which formed clusters of less than five were dropped from the analysis. Thus the present study accounts for 82 of the 101 possible wrong answers. The resulting 14 profile scores (including the two right answer scales) were correlated and the resulting correlation matrix was rearranged into the Simplex order as defined by Guttman (1944).

These clusters were also classified using the reported reasons for selection, so that if a hierarchy emerged, its developmental properties could be examined.

### *Results*

The first question considered was whether or not a hierarchical ordering was evident. Table 1 gives the results of this analysis. The categories are coded by age (10y stands for 10 years old). If more than one subscore set was present for a particular age, these were assigned an arbitrary number (8y4 stands for the fourth subset among the 8-year-olds) to distinguish among them.

It is evident from the arrangement of subtest scores in Table 1 that a clearly definable age-dependent hierarchy was found among these data. In fact, not a single age order displacement occurs. The majority of correlations are significant ( $p < .05$  for  $r = .083$ ) and these clearly support the hierarchy. Most of the significant correlations are low, once again providing evidence nonsupportive of the "linear dependency" assumption.

The reasons given within each cluster for the selection of the particular alternatives in it tended to be logically equivalent between 50 and 60% of the reasons reported.

Proceeding along the hierarchy as given by the Simplex order from low to high, the interpretation given in Table 2 emerged.

TABLE 1  
SIMPLEX<sup>a</sup> ORDER OF THE CORRELATIONAL ANALYSIS  
OF ALL VARIABLES

Right Answer Scales	Wrong Answer Scales	14y+	13y	12y	Bimodal	11y	10y	9y2	9y1	8y3	8y1	Concrete	8y2	8y4
ABSTRACT		<u>.25</u> <sup>c</sup>	.06	.00	<u>-.20</u>	.05	<u>-.22</u>	<u>-.25</u>	<u>-.44</u>	<u>-.46</u>	<u>-.48</u>	<u>-.73</u>	<u>-.60</u>	<u>-.31</u>
	14y+ <sup>b</sup>		.07	.01	<u>-.01</u>	<u>-.09</u>	<u>-.16</u>	<u>-.11</u>	<u>-.18</u>	<u>-.14</u>	<u>-.16</u>	<u>-.22</u>	<u>-.15</u>	<u>-.16</u>
	13y			<u>-.07</u>	.04	.05	<u>-.01</u>	<u>-.08</u>	<u>-.10</u>	<u>-.10</u>	<u>-.16</u>	<u>-.18</u>	<u>-.14</u>	<u>-.05</u>
	12y				.05	.06	.01	<u>-.16</u>	<u>-.08</u>	<u>-.05</u>	.00	<u>-.02</u>	<u>-.06</u>	<u>-.04</u>
	Bimodal					.01	.09	.02	<u>-.00</u>	<u>-.01</u>	.02	<u>-.04</u>	.04	<u>-.05</u>
	11y						<u>.14</u>	.06	<u>-.08</u>	<u>-.10</u>	<u>-.24</u>	<u>-.27</u>	<u>-.24</u>	<u>-.07</u>
	10y							<u>.17</u>	.03	<u>-.01</u>	<u>-.09</u>	<u>-.10</u>	<u>-.01</u>	<u>-.08</u>
	9y2								.05	<u>-.06</u>	.06	.06	<u>.12</u>	<u>-.02</u>
	9y1									<u>.15</u>	.02	<u>.16</u>	<u>.23</u>	.04
	8y3										<u>.36</u>	<u>-.14</u>	<u>.26</u>	.06
	8y1											<u>.37</u>	<u>.26</u>	<u>.22</u>
	CONCRETE												<u>.39</u>	<u>.21</u>
	8y2													<u>.16</u>

<sup>a</sup> The simplex arrangement involves ordering correlations in increasing magnitude vertically upward and horizontally to the left.  
<sup>b</sup> To give a reasonable group size, children of 14, 15, and 16 years of age were combined into one group.  
<sup>c</sup> Underlined correlations are significant at  $p \leq .05$ .

TABLE 2  
INTERPRETATION OF WRONG ANSWER CLUSTERS

General Stages	Wrong Answer Scales	Interpretation	No. of Alternatives in Scale
Personalized interpretations	8y4	Partial translations	5
	8y2	Isolated responses	15
	8y1	Redefining terms	9
	8y3	Personalized answers	7
Literalized interpretations	9y1	Word associations	8
	9y2	Literal reductions	5
	10y	Literalizations	9
Figurative interpretations	11y	Oversimplifications	6
	Bimodal	Irrelevancies	10
	12y	Transpositions	4
	13y	Overgeneralization	6
	14y+	Simplifications	7

The names for these wrong answer categories were assigned as a result of the interpretive analysis conducted in the study.

These 12 classes of response use 82 of the 101 possible wrong responses. There are 19 answers classified as concrete correct answers and 40 as



abstract correct answers. Thus this procedure uses 141 of the 160 (4 × 40) alternatives on the test, or 88% of all possible answers.

Three general classes of wrong answer emerged. First, personalized interpretations all involved the typical egocentric behavior described by Piaget in which the child responds within the framework of her or his own limited personal experience and is unable to entertain externalized reference points. Second, literalized interpretations all involve external reference to the denotive rather than the connotive meanings of the words used. Third, the child in this stage shows greater or lesser facility in the use of the connotative or figurative meanings of words.

A detailed description of the basis for the interpretation of the types of error identified is shown in Table 3. These are given in the same order as in Table 2. Each description contains (1) the name assigned to this type of response by the present researcher; (2) a behaviorally based definition for this type of response derived from the reasons reported by the examinees within each type; (3) an illustration of the behavior with a typical reason

TABLE 3(a)  
TYPES OF ERROR IDENTIFIED IN STUDY  
PART I: PERSONALIZED INTERPRETATIONS

Wrong Answer Scale	8y4	8y2	8y1	8y3
<u>Classification</u>				
Interpretation Assigned	Partial translation	Isolated responses	Redefining terms	Personalization
Behavioral Definition	Only part of the proverb is translated into personal experience.	Explanation of answer selection related to distractor only.	The meaning of the proverb is changed by using alternate meanings for key words.	The child relates both the stem and the foil to the same personal experience, often correctly.
<u>Illustration</u>				
Proverb	THE HOT COAL BURNS, THE COLD ONE BLACKENS	THERE'S MANY A SLIP TWIXT (between) THE CUP AND THE LIP	GOLD GOES IN AT ANY GATE EXCEPT HEAVEN'S	DON'T THROW GOOD MONEY AFTER BAD
Translation Chosen	Leave dangerous things alone.	Don't talk too much while eating.	Anyone would take money.	Don't waste money.
Reason Given	You might get burned badly.	You'll bother people if you talk too much.	I think people should work for their MONEY. (i.e., take = steal).	Because if you waste all your money you will end up poor.
<u>Overall Consistency (Per cent)</u>				
	54.9	56.7	59.4	50.9

TABLE 3(b)  
TYPES OF ERROR IDENTIFIED IN STUDY  
PART II: LITERALIZED INTERPRETATIONS

Wrong Answer Scale	9y1	9y2	10y
<u>Classification</u>			
Interpretation Assigned	Word associations	Literal reduction	Literalization
Behavioral Definition	The child makes a correct association between a word or two in the stem and in the foil and ignores the rest of the problem.	The child extracts most but not all of the literal meaning of the basic communication.	The child translates both stem and foils into a literal statement in his/her own words and matches these literal statements for his answer.
<u>Illustration</u>			
Proverb	THE MORE COST, THE MORE HONOR	DON'T SWAP (TRADE) HORSES WHEN CROSSING A STREAM	ALL IS NOT GOLD THAT GLITTERS
Translation Chosen	Good things have to be paid for in some way.	Don't try some- thing until you're able to do it.	Some things may fool you.
Reason Given	Good things cost more.	You might fall in.	A lot of stones look alike.
<u>Overall Consistency</u> (Per Cent)			
	53.7	57.8	57.1

reported; and (4) the proportion of all respondents in this type whose reported reasons are consistent with the behavioral description. Table 3 emphasizes the three stages in development uncovered among these data in this study.

Problems of order occur at both ends of this sequence. Since the study was conducted in May, the 8-year-old group represents the youngest grade 3 students, about half of whom had become 9 years old during the year. With no reference point lower, it is difficult to establish the order in which the four 8-year-old subscores should be reported. Similarly, since the 14-, 15-, and 16-year-olds are the oldest grade 8 students, the manner in which Oversimplification (with the 11-year-olds) and Simplification (with 14-year-olds) is different is not entirely clear.

It is possible, however, to produce a general two-way classification for these wrong answers. This system is presented in Table 4 and presents the suggestion that particular error types may recur within the general framework of the modifications in orientation that the different stages seem to imply.

When the classified sequence found in this study is rank order correlated with the equivalent categories in the hierarchy found elsewhere among mature adults (Powell, 1970), the correlation is .86. It is evident from this



TABLE 3(c)  
TYPES OF ERROR IDENTIFIED IN STUDY  
PART III: FIGURATIVE INTERPRETATIONS

Wrong Answer Scale	lly	Bimodal	12y	13y	14y+
Classification					
Interpretation Assigned	Over-simplification	Irrelevancy	Transposition	Over-generalization	Simplification
Behavioral Definition	The child correctly translates most, but not all, of the figurative meaning.	The child chooses a true statement which is unrelated to the generally accepted meaning of the passage.	The child changes substantially the figurative meaning of the proverbs.	The child chooses an answer which goes beyond the scope of the usual figurative meaning.	The child correctly interprets most but not all of the usual figurative meaning of the passage. To distinguish this class of responses from the lly one, it is necessary to think in terms of more abstract and/or nearer correctness for this class.
Illustration					
Proverb	RICHES SERVE A WISE MAN BUT COMMAND A FOOL	IT NEVER RAINS BUT IT POURS	THE GRASS IS ALWAYS GREENER IN THE OTHER FELLOW'S YARD	TOO MANY COOKS SPOIL THE BROTH	SPEECH IS THE PICTURE OF THE MIND
Translation Chosen	Don't let money go to your head.	A little is as bad as a lot.	Don't stay in one place to do something.	You'll have too many ideas.	Words paint pictures in your mind.
Reason Given	Money can't (shouldn't) run your life.	If you do something wrong, it's still wrong, even if you do just a little thing or something real bad.	Because if you stay in one place all the time you'll never know what the other place is like.	If you have too many people, they won't be able to decide what to do.	When you're reading, sometimes you can picture what the author's telling you.
Overall Consistency (Per Cent)					
	57.1	59.4	52.9	53.7	62.7

TABLE 4  
DEVELOPMENTAL SEQUENCE OF WRONG ANSWERS

General Strategy	Level of Abstraction of Interpretation							
	Personalized	Literalized	Figurative					
Fragmentation	Isolated responses	Word associations	Oversimplification					
Reduction	Partial translation	Literal reductions	Simplification					
Rearrangement	Redefined terms	Transpositions	(Inversions)					
Extension	(Continuations)	(Extrapolations)	Overgeneralizations					
Transition	[Concrete answers]	Irrelevancies						
Balancing	Personalities	Literalizations	[Abstract answers]					
Age Scale	7	8	9	10	11	12	13	14

( ) Not found in this study.  
[ ] Correct answer subtests.

observation that there is a high level of consistency in the findings between these two independent studies which used different instruments and different age groups. There was also observed to be a perfect agreement for equivalent classes with the error sequence reported in the manual, *Raven's Coloured Progressive Matrices* (1956). Thus this sequence would seem largely to recapitulate the findings of other independent studies on different populations.

### Conclusions

The major characteristics of these newly observed classes of behaviors are fairly evident from the findings reported in this study. Children seem to move from a personalizing stage where they relate all problems to their personal experience through a stage when they take meanings literally before they begin to deal figuratively with meaning in any systematic manner.

It is therefore reasonable to report that this study has discovered a previously unreported aspect of the developmental sequence of behaviors in that wrong answers seem to parallel the observations made by Piaget and Vygotsky, as reported by Stewin and Martin (1974) among others, with respect to the general learning sequence.

Since this study was cross-sectional rather than longitudinal, there is no way of telling if all children go through all these phases. Hence the apparent recurring cycle reported in Tables 3 and 4 may not actually exist. If these recurrences do exist, it may imply that there may be specific thinking styles which particular children characteristically use. The possibility would explain the predictive power of wrong answers observed elsewhere (Powell, 1970). Also, there may be particular transitions through which a child passes and optimal progress takes place.

In any case, the discovery that a type of behavior not usually considered in research studies (namely wrong answers) follows a similar developmental sequence to the one commonly observed in other phases and stage-oriented research has interesting implications for further research. The principal observation here suggests that the way a person interprets and attacks a problem determines the answer selected. In this case, learning may not be a simple linear accumulation of information and skills as suggested by Gagné (1965).

### References

- Bloom, B. S. *Taxonomy of educational objectives: Handbook I: Cognitive domain*. New York: David MacKay, 1956..
- Gagné, R. M. *The conditions of learning*. New York: Holt, Rinehart and Winston, 1965.
- Gorham, D. R. *The proverbs test*. Missoula, Montana: Psychological Test Specialists, 1957.
- Guttman, L. A basis for scaling qualitative data. *American Sociological Review*, 1944, 9, 139-150.
- Powell, J. C. The interpretation of wrong answers from a multiple choice test. *Educational and Psychological Measurement*, 1968, 28(2), 403-412.
- Powell, J. C. *Achievement information from wrong answers*. (Short title). Unpublished doctoral dissertation, The University of Alberta, Edmonton, Alberta, May, 1970.



- Powell, J. C., & Isbister, A. B. A comparison between right and wrong answers on a multiple choice test. *Educational and Psychological Measurement*, 1974, 34(3), 499-509.
- Raven, J. C. *Guide to using the Coloured Progressive Matrices*. London: H. W. Lewis, 1956. (Reference Printing of 1973 pages 21 and following.)
- Stewin, L. L., and Martin, J. The developmental stages of L. S. Vygotsky and J. Piaget: A comparison. *Alberta Journal of Educational Research*, 1974, 20(4), 348-362.

J . K I R M A N

*The University of Alberta*

## The Use of Infra-red False Color Satellite Images By Grades 3, 4, and 5 Pupils and Teachers

*Landsat images are pictures of the surface of the earth produced by satellites. The timeliness and accuracy of these images, colloquially referred to as maps, are of value to social studies teachers. Until now, no research had been undertaken to determine if elementary level children could work with these images, or if elementary level teachers could teach with them. This study is an exploratory attempt to generate preliminary data with a limited sample of participants to determine the feasibility of further definitive research on the topic—in this case, to make an initial judgment whether or not elementary school children and their teachers in this project can deal with infra-red false color Landsat images. The tentative conclusions are that the teachers in this project are able to teach with infra-red Landsat images and children in the project are able to learn about infra-red Landsat images and derive information from them. (Dr. Kirman is Professor of Social Studies in the Department of Elementary Education, The University of Alberta.)*

Today, through the services of space technology, maps of the earth's surface are made by orbiting satellites. These maps, known as Landsat and Seasat images (land satellite and sea satellite images respectively) and formerly known as ERTS (Earth Resources Technology Satellite) images, were first obtained when ERTS-1 was launched, July 23, 1972. A second such satellite was launched January 23, 1975. These satellites follow each other on a nine-day interval (Bricker & Campbell, 1975). These satellites are in a near polar orbit, 560 miles high, each covering almost all of the earth's surface every eighteen days (National Aeronautics and Space Administration, 1973 [NASA]). Each satellite picture contains an area of 13,000 square miles, resolving features not smaller than 30 meters (NASA, 1973).

The accuracy and timeliness of these images, as well as their uses are of extreme value:

Agricultural fields, surface water, types of land forms, patterns of urban



development and other changes in the earth's surface from natural or man-made causes are now observed on a regular basis . . . . ERTS-1 also carries a Data Collection System (DCS) that acquires water quality, rainfall, snow depth and seismic activity information from remotely located sites in North America. (NASA, 1973, p. 2)

The scale of a Landsat image is 1:1,000,000; however, these images can be enlarged without loss of detail up to a scale of 1:250,000. Several images are produced in various visible and nonvisible light spectrum bands both in black and white and in composite false color (NASA).

These images are perhaps more informative and accurate than any hand drawn map based upon previous cartographic technology. Among possible uses of the images are: comparative growth of communities, land use, nature of vegetation, evidence of pollution, examination of geologic patterns.

With a device called a density slicer, even more uses are presented, from the identification of specific mineral and vegetable surface features to the quality of the surface soil (Bricker & Campbell, 1975). However, this latter item is a most expensive and sophisticated tool, presently out of financial reach of almost all elementary and secondary school systems.

Given the nature of these Landsat images and the potential for classroom instruction that might evolve, the following questions were raised:

1. Can elementary level children obtain data from Landsat images?
2. How would elementary teachers use Landsat images, assuming children on these grade levels are capable of dealing with Landsat images?

An examination of the literature shows no studies of the use of Landsat images with elementary level children.

B. E. Kingston (1969) used aerial photographs with children in grades one, two, and three. The maps used on each grade level were to a maximum scale of 1:31,800. Kingston found that these children could read vertical photographs, but had difficulty in interpreting them without previous teaching.

K. G. Dueck (1969), in a study similar to Kingston's utilizing aerial photographs of the same scale with children in grades four, five, and six, found that these children could read aerial photographs; there was no problem with scale; interpretation of these maps was not correlated with map reading ability and age; and that grade 5 performed better than grade 4 or 6.

Concerning the capacity of elementary level children to deal with maps, Brown and others (1970), in a study in English schools, examined boys and girls aged eight to fifteen years old. Some suggested results infer difficulty for younger children in map reading (Descriptor Paragraph).

In a study by Blaut (1971) of children aged three to twelve dealing with "mapping and free environmental behavior" among other items, mapping behavior was found to be developed in preschool children.

The Kingston, Dueck, and Blaut studies appear to confirm that elementary level children are capable of dealing with aspects of aerial photographs and maps. While the Brown study does not specifically deal with aerial photographs, it appears to suggest difficulty in the comprehension of certain aspects of maps by elementary level children. However, items such as the types of maps, the nature of the research, socioeconomic levels,

as well as the nature of the English social studies curricula in comparison with those of North America may also have some bearing on the differences between this study and the others.

### *Definitions*

**Landsat Image:** a representation of the land surface of the earth produced with the aid of an orbiting satellite, often referred to as a map, both terms being interchangeably used by teachers who find “map” a more familiar term at this time.

**Infra-red False Color:** that portion of the spectrum corresponding to infra-red, and used to delineate vegetation (green) on Landsat images.

**Landsat:** Land satellites, scanning the land surfaces of the earth.

**Seasat:** Sea satellites, scanning the ocean surfaces of the earth.

**ERTS:** Earth Resources Technology Satellite. Formerly the name of Landsat satellites.

**Visible Light Spectrum;** that portion of the light spectrum visible to the human eye.

**Composite False Color:** color produced on Landsat images composed of data derived from different satellite spectrum scans of the same area.

**Non Visible Light Spectrum:** that portion of the light spectrum not visible to the human eye.

**Polar Orbit:** the orbit of a satellite where the path is over the earth’s poles.

**Vertical Aerial Photographs:** photographs taken in an aircraft of the surface of the earth directly below the aircraft.

### *Rationale*

The rationale for undertaking this study with elementary level children is that Landsat images are a new social studies resource, and the prime researcher was curious as to whether or not it could be used by young children.

In addition, while it might be presumed that secondary level students would probably be able to cope with many aspects of Landsat images, no such presumption can be freely made for the elementary level students, given the magnitude of the scale involved. Hence, if such relatively young children could cope with these new images, the potential for early introduction of them might be discerned.

### *Method*

#### *Instruments*

Three instruments were used in this exploratory project: teacher-made tests for each class; an “Interview Questions for Children on Landsat Maps” administered and prepared by the prime researcher; teacher questionnaire prepared by the prime researcher.

The instruments are subject to the following:

1. No validated instrument has been found that can test the abilities of elementary level school children to read infra-red false color Landsat images.



2. No previous studies have been found utilizing Landsat images with elementary level school children in order to provide a comparative study to extrapolate testing tools for this study.
3. The purposes of the study are to determine if the participating elementary level school children are capable of deriving information from infra-red false color Landsat images, and if their participating teachers can teach about them.

Thus, the teacher-made instruments would necessarily be:

1. Tentative rather than definitive;
2. Specific to this study, rather than general;
3. Designed to elicit a response unique to the teaching procedures of each class and the information taught;
4. Designed to elicit responses to Landsat image questions considered correct by each teacher.

Under these circumstances, the teacher-made instruments may be considered valid for their particular classes. Since these instruments are unique to these teachers and children and are not to be used again, the element of reliability appears not to enter into this consideration.

The "Interview Questions for Children on Landsat Maps," administered by the prime researcher, was used to elicit responses to questions based specifically on Landsat images. The object was to determine if children of varying abilities, using infra-red Landsat images different from those used in their class, were able to derive information from them and answer questions about them.

This instrument was designed specifically to generate tentative data on the ability of a select number of children of varying abilities to answer specific questions about infra-red Landsat images (vegetation and snow cover images). Since the questions were specific to the Landsat images in question and since the children appeared able to understand the questions and respond to them, it may be tentatively considered that this tool was valid for the purposes stated.

Since the questions are such that they may be applied to any infra-red false color Landsat images, the instrument may be tentatively considered as reliable for the purpose stated.

### *Research Questions*

In order to determine if children on the grade 3, 4, and 5 level were capable of dealing with Landsat images, the following questions were formulated:

1. Can grade three, four and five teachers work with Landsat images in their planning for classroom activities?
2. Can grade three, four and five children understand what Landsat images are by deriving information from them?
3. Will grade three, four and five teachers be able to teach their students about Landsat images?
4. Will grade three, four and five teachers have to modify Landsat images in order to use them with their students?

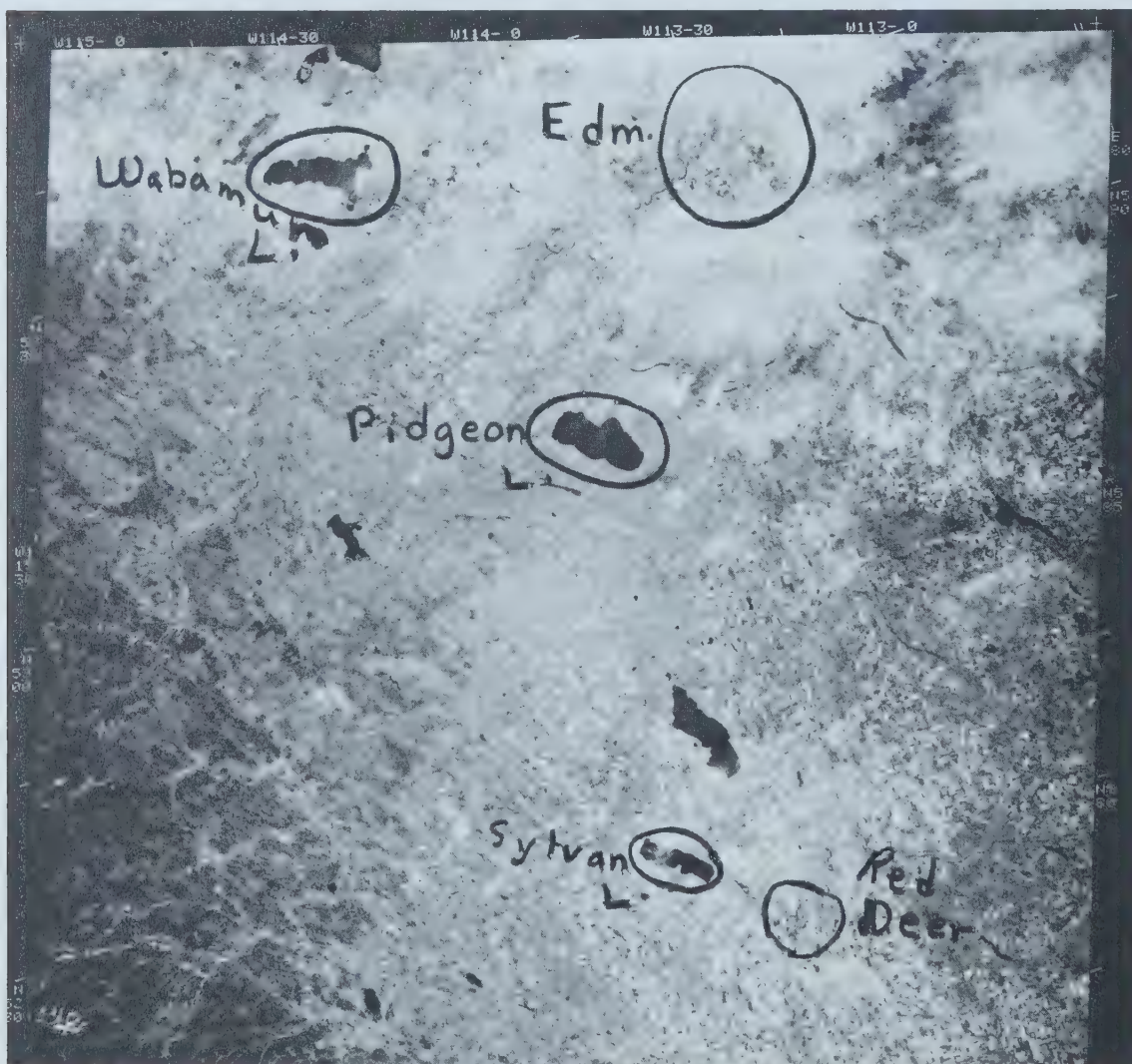
### Subjects

The rationale for the use of Landsat images with grades three, four and five is that the Alberta social studies curriculum calls for community studies in grade three, Alberta in comparison with other regions of the world in grade four, and regional studies of Canada in grade five. These topics can lend themselves to the use of maps for classroom activities.

In cooperation with the Edmonton Separate School Board, one teacher and class on each grade level were selected by the school board's social studies supervisor on the basis of interest and ability. The teachers and students had no previous exposure to Landsat images. Prior to the classroom phase of this exploratory project, the teachers and research director met twice as a group. The first meeting was at the Alberta Remote Sensing Center in Edmonton for approximately one hour. There, they were instructed in interpreting Landsat images and in the technology involved in obtaining the image.

A second meeting was held for approximately one hour at the Separate School Board headquarters for a brain storming session on possible ways of

FIGURE 1  
TYPICAL FALL LANDSAT IMAGE WITH GRADE 3 STUDENT NOTATIONS.





using the Landsat images in class. Teachers also received several National Aeronautics and Space Administration publications on Landsat images.

At the first meeting, it was determined that two Landsat images would be used: one with snow cover and one without. The teachers decided that infra-red false color might prove more effective with the children due to the attractiveness of the color and the apparently greater contrast between earth surface features than on any of the black and white images. Landsat images are not made in true color.

### *Classroom Materials*

The images selected by the teachers were of the Edmonton Region, one was a fall image without snow cover; the other was a spring image with snow cover. Both images were of approximately the same area, with the spring image covering a more southwesterly area than the fall image. The size of the images was 18.2 cm  $\times$  18.4 cm each. Figure 1 shows a typical image complete with third-grader's notations.

Enough prints of the images were provided for two children to work on a single set. In addition, some magnifying glasses were provided. Each teacher had available an overhead projection size transparency of each image in infra-red false color, a class set of standard maps that included the area of the Landsat images, and an overhead projection transparency of the standard map with the approximate areas of the Landsat images marked on them. Black China markers were also provided for the children to mark the surface of the images.

### *Procedure*

Since there were no studies on the use of these types of images with elementary school children, the teachers were asked to teach about them in the way they felt was best for their classes. The classes were heterogeneously grouped as shown in Table 1.

In order to allow for complete flexibility of class instruction and to avoid teaching to a test, no uniform written posttest of all the children was given. The teachers were requested to prepare an examination and evaluate their students following the classroom instruction phase, and to determine prior exposure to these types of images.

Following the classroom evaluation, the prime researcher orally interviewed three children on each grade level. Selected by the teachers, these children were considered above average, average, and below average. The teachers also received a questionnaire on their participation in the project. All data were submitted to an external examiner for review.

### *Classroom Phase*

Prior to starting the instruction phase of the project, the teachers were asked to check the medical records to ascertain if any of the children had vision problems and to specially observe those who did.

Since there was no experience base concerning Landsat images on the elementary level, the teachers emphasized different aspects of them. Thus, the grades three and four teachers emphasized radiation that allows images

TABLE 1  
SUBJECT DATA

		Mean	S.D.
<u>Grade 3</u> (N=25: Male, 15; Female, 10)			
Reading Level		4.07	.800
I.Q. (Primary Mental Abilities Test)		108.5	9.88
<u>Grade 4</u> (N=25: Male, 16; Female, 9)			
Reading Level		4.93	1.21
I.Q. (Lorge Thorndike)	Verbal	106	13.02
	Nonverbal	109.70	14.01
<u>Grade 5</u> (N=19: Male, 12; Female, 7)			
Reading Level		4.42	1.00
I.Q. (Lorge Thorndike)	Verbal	91.31	13.91
	Nonverbal	87.25	14.19

to be made while the grade five teacher emphasized specific geographic features.

All teachers provided some background information on how the images were produced. The grade three and four teachers prepared some ditto sheets of National Aeronautics and Space Administration materials for distribution to the children, and had the students prepare a special folder for the Landsat activities. It may be noted that the grade three and four teachers were both at the same school. There was virtually no prior exposure of the students to satellite images, as determined by the teachers.

During the course of the project, the research director visited all classrooms a minimum of three times. Teachers were prepared and the students participated in an interested orderly manner. All classes were well organized. When not conducting a question and answer activity, the teachers moved among the students providing advice and guidance.

Class activities consisted of: basic factual instruction; questions and application of information to the images at hand; examining a single Landsat image first, usually the fall one, for surface features; examining the second image for these surface features; checking for differences between the fall and spring image; comparing the Landsat images with standard road maps (grade three did not use these maps, but examined the overhead projector transparency of the road map—there was a road map displayed on the back wall of the classroom during the entire project).

Teachers began working with the fall image first. Then when the children were able to work with it, the spring one was introduced. After examining the spring Landsat image and becoming familiar with it, the children were given both images to be used for comparative purposes. Table 2 indicates class variation between teaching time, period length, and number of instruction periods.



TABLE 2  
TEACHING TIME, PERIOD LENGTH AND NUMBER

Grade	Hours of Teaching	Period Length	Instruction Periods
3	5	45 minutes	6 - consecutive with a double period later in the day, if necessary
4	4½	45 minutes	6 - consecutive
5	5½	30 minutes	11 - over a period of 16 days, with one lesson on any one day.

One aspect of the instruction was that the children were able to write and mark the surface of the Landsat images with easily erasable China markers. When the overhead transparency projector was used, children were encouraged to come to the screen and point out features being discussed or newly discovered.

In all classes, children worked in teams of two. The reason for this was to minimize expenses for the maps. A second benefit was that each child was able to interact with at least one other child.

### *Results*

#### *Teacher Evaluations*

The teacher-made examinations differed greatly concerning content, number of questions, and manner of administration. However, specific questions in almost all tests could be designated either as interpreting the images or applying the images to a problem. In this way, a very rough measure could be obtained of the number of students able to interpret the images or apply the images to a problem. Table 3 shows that a large number of children in each grade appeared capable of interpreting some aspects of Landsat maps and that a large number of children in grades three and four appeared capable of applying Landsat maps to a problem.

Class scores varied, as shown in Table 4. However, by inspection the scores show that the children were able to cope with the examination questions and while some did have relatively lower grades than others in their classes, many scored quite high. In general, it appeared that the grade range roughly suggests the ability to deal with questions based on Landsat maps prepared by the teachers.

#### *Interview Evaluation*

Following the classroom phase of the project, the prime researcher interviewed three children on each grade level: above average, average, below average for the class as determined by their teachers. Each child was asked a series of 15 questions, 13 of which had one or more responses that could be graded as correct or incorrect. Based upon these 13 questions, there was a possible total of 20 correct answers.

The questions were based upon a fall and a spring Landsat image, and one question was based upon the Landsat images and a road map of the area

TABLE 3  
DISCRIMINATING QUESTIONS FOR INTERPRETATION AND  
APPLICATION OF LANDSAT IMAGES

<u>Interpretation</u>			
<u>Grade 3</u>			
Question 6.	What types of land surface are seen on these maps? Name them.		
	Full Credit 14	Half Credit 9	Total Credit 23
<u>Grade 4</u>			
Question 5.	What is the main thing to be seen when you look at any Landsat map?		
			Total Credit 13
<u>Grade 5</u>			
Question 5.	(one among a series of interpretive questions) On this summer map, the large blue areas represent: (a) rivers (b) lakes (c) forests (d) mountains.		
			Total Credit 19
<u>Application</u>			
<u>Grade 3</u>			
Question 9 (a).	What can these ERTS maps tell farmers?		
	Full credit 14	Half Credit 4	Total Credit 18
(b).	How can the maps help the Forestry Department or the lumberman?		
	Full credit 20	Half Credit 3	Total Credit 23
<u>Grade 4</u>			
Question 6.	Name three different groups who could use Landsat maps for their benefit?		
	Full Credit 8	Partial (2) 5	Partial (1) 3
			Total Credit 16
<u>Grade 5</u>			
Questions were mainly interpretive.			

(Table 5). The images and map were of a different area than those used in the class. Question 10 of the series required the children to pick out a town on the Landsat images. No child was able to do this, due to the small size of the towns in the area, and this question may not constitute a valid measure of the children's capability to do this (Table 6).

Research question 2 regarding grade three, four, and five children's understanding of Landsat images and deriving information from them appears to be answered affirmatively. Both the teacher-made tests and the oral evaluation show that the children were able to derive information from the Landsat images. One notable discrepancy did occur. The mean of the teacher-made test grades were comparatively low for the grade four girls with relatively high I.Q.s and reading scores. In discussing this with the grade four teacher, it was noted that the girls were generally less enthusiastic



TABLE 4  
SCORES FOR THE TEACHER-MADE TESTS

	Mean	S.D.	Range
<u>Grade 3</u>			
Class	78.1	13.73	45-100
Males	78.2	14.76	45-100
Females	78	12.95	65- 90
<u>Grade 4</u>			
Class	54.74	23.18	0-100
Males	57.5	22.61	30-100
Females	43.8	29.25	0- 80
<u>Grade 5</u>			
Class	69.1	12.02	44- 92
Males	69.8	15.20	44- 92
Females	68.6	4.86	60- 76

TABLE 5  
INTERVIEW QUESTIONS FOR CHILDREN ON LANDSAT IMAGES

1. What are these? (Pointing to both Landsat images)
2. How were they made?
3. What are the differences between these two maps?
4. Can you tell if there are any growing plants on these maps? How?
5. Can you find a lake on both of these maps?
6. Can you find a river on both of these maps?
7. Can you find a road on both of these maps?
8. Can you find any clouds on both of these maps?
9. Can you find any farm land on both of these maps?
10. Can you find a town on both of these maps?
11. Here is a regular map of the land shown by the Landsat map. Find the town of High Level on this map. Now find High Level on a Landsat map.
12. What could people use these maps for?
13. Can you tell me anything else you know about these maps?
14. What else would you like to learn about these maps?
15. If you were in the lumber business, could Landsat maps help you? How?

about the map activities than were the boys, as evidenced both by classroom response and the amount of outside materials brought to class.

One item of data that could have been derived from this project was which of the two images was easier to work with, the vegetation cover or the snow cover. This was not done since any data would be misleading. This is

TABLE 6  
DATA OF THE INTERVIEW EXAMINATION  
(A = Above Average B = Average C = Below Average)

Grade	Age	Sex	I.Q.		Reading	Correct Responses	Percent
<u>Grade 3</u>							
A	7.3	M	N/A		4.8	13	.65
B	8.1	F	129		4.7	13	.65
C	8.1	M	109		2.7	13	.65
<u>Grade 4</u>							
			<u>Verbal</u>	<u>Nonverbal</u>			
A	9.4	M	112	131	6.2	12	.60
B	9.1	M	122	118	5.6	17	.85
C	8.11	F	108	115	4.9	5	.25
<u>Grade 5</u>							
A	10.3	F	114	111	5.8	13	.65
B <sup>a</sup>	9.9	F	N/A	N/A	5.3	8	.40
C	11.7	M	N/A	N/A	N.A	11	.55

<sup>a</sup> Teacher noted child as nervous and inattentive with family mental health and marital instability.

due to the variation in observation of specific features because of the intensity of the infra-red false color and contrast with nonvegetation varying from region to region, and whether or not inhabited areas or roads are cleared of snow in the winter. Thus, each region's Landsat images varying from season to season must be examined on their merits. However, it appears that greater contrast between ground cover and items of interest can promote easier identification of these items.

The data seemed consistent with Kingston's (1969) findings that youngsters in grade three can read vertical photographs, and that without previous teaching, there was some difficulty in vertical aerial photograph interpretation.

The data also appear to be consistent with Dueck's (1969) findings that there was no significant correlation between chronological age and map reading ability. Although due to the variation in class ability, experience of the teachers, and the differing approaches taken by the teachers in this project, this consistency may not be valid.

Since there appeared to be no problems concerning map scale, the project data seem consistent with Dueck's similar findings. While Dueck found grade five subjects to perform better than grade four or grade six subjects, the grade three subjects in this project appear to have performed better on their evaluations than the grade four or grade five children. This may possibly not be considered a valid comparison for the reasons noted above.

*Conclusions and Recommendations*

Research questions 1, 2, and 3 appear to be answered affirmatively for the participants in this project. Research question 4 appears to be answered negatively for the participants in this project.



Since children on the grade three, four, and five level seemed to be able to work with infra-red false color Landsat images to obtain data from them, it may be tentatively concluded that such Landsat images can be used on these levels. However, the extent to which children can apply Landsat image derived data to problems remains to be explored.

Teachers of grade three, four, and five children in this project appeared to be able to teach about Landsat images to their students. However, the level of instruction appeared related to the background knowledge of the teacher. It may be concluded that a more extensive program of instruction than was provided in this project is necessary to provide teachers with Landsat image background information.

### *Implications of this Project for Social Studies*

The ability to work with Landsat images provides a new dimension of study. The Landsat images being continually made of the same areas every nine days can provide timely, accurate data on urban and rural development, pollution problems, agricultural conditions, as well as comparative examinations of almost any surface area of the planet.

In such studies, economic, sociological, and political as well as geographic implications of these images can be discussed with visual accuracy. Over a period of time, a historical examination of areas might be undertaken with a chronological series of images.

To reiterate, the important characteristic of Landsat images for social studies is their accuracy and timeliness. Presently, there are no other materials that can do the same or better.

### *Suggestions for Further Research*

1. Is there a sex related difference regarding grade four students' interest and achievement with Landsat images?
2. Can children on grade three, four, and five levels work with the black and white Landsat images? If so, which types?

A more definitive project allowing greater scope for claims of effectiveness of using Landsat images with elementary school children might be undertaken with the following considerations:

1. A larger selection of students on each grade level: a minimum of three classes, or as many as needed to reach 100 students on each grade level.
2. All students to be as nearly as possible within the range of "average." If not, then as intellectually comparable as possible.
3. All students to be as nearly as possible within the same socioeconomic level.
4. All students to have "normal" reading levels. That is, no students with serious reading problems in the sample, if possible.
5. Participating teachers to be as comparable as possible to each other as to professional training, teaching experience, ability, and interest.
6. Participating teachers to follow specific teaching procedures on each grade level.
7. Teaching objectives to be specified before classroom phase is to begin.
8. All children to be pre- and post-tested, using uniform examinations for

all grade levels, and uniform examinations on specific grade levels where special circumstances relating to that grade level merit such examinations.

9. A thorough training program for participating teachers in teaching about Landsat images, and how these images are produced and used. Five hours is suggested for this phase, as a minimum.
10. External evaluation at all phases of the project.

This project was funded by a grant from the Alberta Advisory Committee for Educational Studies. The cooperation of Dr. Donald Massey, University of Alberta, as external examiner, and the Edmonton Separate School Board for its facilities is gratefully acknowledged.

#### *References*

- Blaut, J. M. Studies of place perception in elementary and pre-school education. Vol. I. Final report. Clark University, 1971. (Clearinghouse Accession Number: PS005603. Descriptor.)
- Bricker, C. D. (Coordinator), & Campbell, K. M. (Technologist). Alberta Remote Sensing Center, Edmonton, Alberta. Interview, Nov. 21, 1975.
- Brown, T. X. An investigation into the optimum age at which different types of map questions may best be set to pupils in the teaching of geography. International Geographical Union, 1970. (Clearinghouse Accession Number: So002755. Descriptor.)
- Dueck, D. G. Reading and interpreting vertical aerial photographs in the primary grades (4, 5, 6). Unpublished M.Ed. thesis, University of Calgary, 1969.
- Kingston, B. E. Reading and interpreting vertical aerial photographs in the primary grades (1, 2, 3). Unpublished M.Ed. thesis, University of Calgary, 1969.
- National Aeronautics and Space Administration. *Photography from space*. Goddard Space Flight Center, Maryland: N.A.S.A., 1973.



R. L. NOLL

*State College Area Schools*

D. J. WILLOWER

and

J. J. BARNETTE

*The Pennsylvania State University*

## Teacher Self-Actualization and Pupil Control Ideology-Behavior Consistency

*The hypothesis that teacher level of self-actualization would predict the consistency between teacher pupil control ideology and behavior was tested in this study. Data were gathered on the levels of self-actualization and pupil control ideologies of 84 teachers, and data were collected from students concerning the pupil control behavior of these teachers. The hypothesis was not rejected. Previous research that showed a positive relationship between teacher self-actualization and humanism in pupil control ideology was replicated with similar results. However, teacher self-actualization was not associated with humanism in pupil control behavior. (Dr. Noll is school psychologist, State College Area Schools, State College, Pennsylvania; Dr. Willower is professor of education and Dr. Barnette is assistant professor of education, The Pennsylvania State University.)*

In a previous study (Jury, Willower & DeLacy, 1975) it was found that teacher level of self-actualization was directly related to teacher humanism in pupil control ideology. The present research is an effort to build upon that work. This article reports on an investigation of the relationship of teacher self-actualization and the consistency of teacher pupil control ideology and behavior.

The concept of self-actualization (SA) has been developed most fully by Maslow (1957, 1964, 1968, 1970, 1971). The self-actualizing person is one who, compared with others, makes full use of his capabilities, functions in an autonomous manner, lives in the present, and has a benevolent view of people and of life (Maslow, 1970).

The idea that pupil control is a key feature of the social system of the public school was one outcome of a field study (Willower & Jones, 1967; Willower, 1971) that led to a line of inquiry (Willower, 1975) on educator pupil control ideology (PCI) and pupil control behavior (PCB). In that work, both pupil control ideology and behavior were conceptualized on a humanistic-custodial continuum. Humanism stresses trust of pupils and self-discipline, while custodialism emphasizes distrust of pupils and imposed discipline.

### *Hypotheses*

Three hypotheses were tested in this investigation. The major hypothesis predicted that there would be a direct relationship between teachers' levels of SA and the consistency of teachers' PCI and PCB. The rationale for this hypothesis was quite simple. The individual at a relatively high level of SA should be one who has satisfied basic needs, is positive toward others, and is autonomous and self-directed. Such a person is likely to behave in a manner consistent with his beliefs.

The second hypothesis was a replication of the research by Jury, Willower and DeLacy (1975) noted earlier. It stated that there would be a direct relationship between teachers' levels of SA and humanism in PCI. The third hypothesis tested was similar to the second, but it examined teacher PCB, not PCI. It posited a direct relationship between level of SA and humanism in PCB.

### *Method*

#### *Instruments*

The Personal Orientation Inventory (Shostrom, 1968) was used as a measure of SA. It measures positive health or SA rather than pathology (Shostrom, 1964). The inventory consists of 150 paired items; the higher the score to a theoretical maximum of 150, the greater the level of SA. Test-retest reliabilities above .90 have been reported for this instrument, and it has been shown to differentiate between persons judged by clinical psychologists to be self-actualized and nonself-actualized (Shostrom, 1968).

Teacher beliefs and behavior were measured respectively by the PCI Form and the PCB Form. These devices are companion instruments in that both are based on the same humanistic-custodial continuum. Split-half reliabilities for the PCI Form ranged from .91 to .95, and it was shown to discriminate between teachers and schools judged to be humanistic or custodial (Willower, Eidell & Hoy, 1973). PCB Form reliability as estimated by Cronbach's alpha was .92. Item-scale correlations for the PCB Form averaged .81 and one-way analysis of variance indicated that the instrument differentiated among subjects while clustering within subjects (Helsel & Willower, 1974).

Both the PCI and PCB Forms consist of 20 items. The scoring range on each is from 20 to 100, the higher the score the more custodial the ideology or behavior. While the PCI Form is completed by the subject whose ideology is being tapped, the PCB Form is completed by students with their teacher as the object of response. A given teacher's PCB Form score is the mean of the PCB Form scores of the students in that teacher's class.



### *Sample and Procedure*

The sample consisted of selected teachers and students in a single school district in the northeastern United States. The district's 11 elementary schools, two junior high schools and one senior high school serve a small city and its surrounding suburban-rural areas. The faculty of the school system approximated state averages for salary and years of teaching experience.

It was decided to gather data from all of the school district's fifth and sixth grade teachers at the elementary school level and all of the mathematics and social studies teachers at the secondary school level, as well as from the students of the fifth and sixth grade teachers and the students in one of the classes of the secondary school teachers. Fifth and sixth grade teachers were chosen since students at those grade levels would have no difficulty responding to the PCB Forms, and mathematics and social studies teachers were chosen because they represented major, but quite different, instructional areas. There were 49 fifth and sixth grade teachers in the district, and 47 participated in the study; there were 46 mathematics and social studies teachers, and 38 participated. A total of 2,068 students took part in the study.

Teachers were given the PCI Form and the Personal Orientation Inventory at grade or departmental meetings, and a date was set for the collection of these forms and the administration of the PCB Forms to students. However, a specific time or class period was not designated. This meant that the researcher's schedule rather than the teachers' choices dictated which of a given secondary school teacher's classes completed the PCB Forms on that teacher. However, this was a convenience procedure rather than a random one.

The response of one elementary teacher and of a number of students were incomplete; thus, the final sample used in the analyses consisted of 84 teachers, 46 at the elementary level and 38 at the secondary, and 2,044 students, an average of 24.3 per teacher.

### *Results*

The major hypothesis tested stated that there would be a direct relationship between teachers' levels of SA and the consistency of teachers' PCI and PCB. The dependent variable, PCI-PCB consistency, was the absolute difference in PCI and PCB scores calculated after PCI and PCB raw scores had been converted to standard scores. Data on the several variables examined in the study are presented in Table 1.

The Pearson product-moment correlation coefficient between the teachers' SA scores and their difference scores was  $-.225$ . The minus sign was an artifact of the way consistency was measured; as SA scores increased, difference scores decreased. With 82 degrees of freedom, the correlation was significant at the .05 level. Hence, the major hypothesis was not rejected.

The second hypothesis predicted a direct relationship between teachers' levels of SA and humanism in PCI. The correlation of  $-.486$  was significant beyond the .001 probability level, indicating that the hypothesis could not be

TABLE 1  
TEACHER SA AND PUPIL CONTROL DATA ( $N = 84$ )

Variable	Score	SD
SA	102.1	10.9
PCI	52.2	8.8
PCB	48.4	9.6
PCI-PCB Standard score difference	10.4	7.1

rejected and that the earlier work of Jury, Willower and DeLacy (1975) was supported in this investigation.

The third hypothesis proposed a direct relationship between teachers' levels of SA and humanism in PCB. In this case the  $r$  of  $-.060$  was not significant and the hypothesis was rejected.

In addition to the tests of these hypotheses, a number of other analyses were carried out. They included tests of the three hypotheses in subsamples of teachers, and comparisons of various mean scores for several categories of teachers. These are reported in detail in Noll (1976). Only those bearing on the hypotheses are noted here.

When the major hypothesis was tested separately within the teacher subsamples based on grade level, sex, and years of experience, it held up only in the case of teachers having more than five years' experience. For these 58 teachers, the correlation of  $-.278$  was significant at the  $.05$  level. In the case of the second hypothesis, the relationship between teacher SA and PCI held for every category examined. The correlations ranged from  $-.456$  for elementary teachers to  $-.541$  for the 33 male teachers in the sample. In contrast, none of the correlations between teacher SA and PCB were significant. In fact, the highest of these correlations,  $.287$  for the 26 teachers having five years' or less experience, although not significant, was in the opposite direction of the prediction.

### *Discussion*

Before discussing our findings, we wish to make the usual disclaimers. Although results were significant in several cases, correlations were moderate and the variance accounted for was limited. In other words, much of the variance in the several dependent variables is associated with unknown variables other than self-actualization. Also, the research utilized paper and pencil measures throughout with all the limitations of that medium. Finally, the sample was not a large one, a consideration that became more material when subsamples were examined.

Having expressed these cautions, we will take our findings at face value in order to offer some speculations concerning them. For the sample studied, there was a clear cut relationship between teacher SA and PCI, a modest relationship between teacher SA and PCI-PCB congruence, and no relationship between teacher SA and PCB.



The relationship between SA and PCI is one between orientations that can be considered internal to the individual. They are internal in the sense that there is no *necessary* connection between the orientations and behavior. That the more self-actualized teachers were more humanistic in PCI both in this and a previous study indicates complementary orientations. A self-actualized person tends to see people and the environment as relatively benign, and the self-actualized teacher is likely to view students as trustworthy and favor a humanistic control ideology. That internal orientations of this kind should complement one another makes sound theoretical sense.

When the major hypothesis guiding this investigation was tested, we found that more self-actualized teachers tended to be more consistent in their beliefs and behavior, specifically in terms of their beliefs and behavior regarding pupil control. This suggests that persons having higher levels of SA are less vulnerable to the organizational, peer group, and other pressures that commonly intervene to divorce beliefs and behavior, shaping behavior to fit situational demands. However, the strength of the relationship found between teacher level of SA and PCI-PCB consistency was only moderate. This could be interpreted to mean that the teachers in our sample adapted to the classroom environment by basing much of their pupil control behavior on situational pressures and particular circumstances, but that the more self-actualized teachers managed to maintain at least some consistency between beliefs and behavior.

The fact that there was no relationship between teachers' levels of SA and their PCB suggests adaptive behavior more strongly than did the modest results of the test of the major hypothesis. We conjecture that teachers' PCB is influenced to a degree by their attitudes, to a degree by organizational and social system demands, and to a degree by the realities of the work of teaching. That work requires that all teachers respond to such classroom realities as student questions, requests, and misbehavior, as well as manage a wide range of activities of a logistical kind. And we have not even mentioned the instructional aspect of teaching. It would seem that the setting within which teachers work is one characterized by a high level of demands, demands which markedly reduce leeway for the expression of teacher behavior. In any case, the interrelationships between personality, beliefs, behavior, and the demands of the organization, the social setting and the work itself are subtle and complex.

In summary, it was found that teachers who are more self-actualized tend to be more humanistic in PCI than their less self-actualized counterparts, and it was argued that this relationship stems in part at least from the association of two complementary orientations internal to the individual. It was found that teachers who are more self-actualized are more likely than their less self-actualized colleagues to be consistent in PCI and PCB. The speculation was advanced that the modest correlation that characterized this relationship meant that numerous forces effectively intervene to channel behavior in other than the direction dictated by one's beliefs, but that more self-actualized persons are somewhat less vulnerable to these influences than less self-actualized persons. Finally, no relationship was found between teacher SA and PCB, and it was conjectured that this indicates the extent to which social and organizational forces and the

character of the work of teaching itself impose demands upon behavior which even quite different individuals find compelling.

#### References

- Helsel, A. R., & Willower, D. J. Toward definition and measurement of pupil control behavior. *Educational Administration Quarterly*, 1974, 12, 114-123.
- Jury, L. E., Willower, D. J., & DeLacy, W. J. Teacher self-actualization and pupil control ideology. *Alberta Journal of Educational Research*, 1975, 21, 295-301.
- Maslow, A. H. Self-actualization and beyond. In J. F. T. Bugental (Ed.) *Challenges of humanistic psychology*. New York: McGraw-Hill, 1957.
- Maslow, A. H. *Religions, values, and peak experiences*. Columbus: Ohio State University Press, 1964.
- Maslow, A. H. *Toward a psychology of being* (2d ed.). New York: Van Nostrand Reinhold, 1968.
- Maslow, A. H. *Motivation and personality* (2d ed.), New York: Harper and Row, 1970.
- Maslow, A. H. *The farther reaches of human nature*. New York: Viking Press, 1971.
- Noll, R. L. Teacher self-actualization and pupil control ideology-behavior congruence. Unpublished doctoral dissertation. The Pennsylvania State University, 1976.
- Shostrom, E. L. An inventory for the measurement of self-actualization. *Educational and Psychological Measurement*, 1964, 24, 207-218.
- Shostrom, E. L. *Manual for the personal orientation inventory: An inventory for the measurement of self-actualization*. San Diego: Educational and Industrial Testing Service, 1968.
- Willower, D. J. The teacher subculture. In L. W. Drabick (Ed.), *Interpreting education: A sociological approach*. New York: Appleton, Century, Crofts, 1971.
- Willower, D. J. Some comments on inquiries on schools and pupil control. *Teachers College Record*, 1975, 77, 219-230.
- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology* (2d ed. with an annotated bibliography of pupil control studies). University Park: Penn State Studies No. 24, 1973.
- Willower, D. J. & Jones, R. G. Control in an educational organization. In J. D. Rath and others (Eds.), *Studying teaching*. Englewood Cliffs, N.J.: Prentice-Hall, 1967.



S . J . H . M c C A N N

London, Ontario

and

R . J . F I S H E R

University of Saskatchewan

## The Interactive Relationships of Teacher Directiveness and Student Authoritarianism and Dogmatism to Grades and Satisfaction

*This study was designed to determine which of the interactions of teacher directiveness and student authoritarianism, teacher directiveness and student dogmatism, teacher authoritarianism and student authoritarianism, and teacher dogmatism and student dogmatism account for unique variance in grades and satisfaction. Six teachers and 174 students in grades 11, 12, and 13 English classes completed the F-Scale and the Dogmatism Scale. Students also completed a teacher satisfaction scale and the SPOTS teacher directiveness scale. Each student's English grade, IQ, sex, English teacher's sex, year, and academic program were determined from school records. Three dependent variables, grades, z-grades, and satisfaction were developed. Perceived teacher directiveness interacted with student authoritarianism in the determination of z-grades ( $p < .05$ ), and with student dogmatism in the determination of satisfaction ( $p < .01$ ). These interaction effects remained significant when teacher directiveness was statistically controlled. Both interactions were related to grades,  $p < .01$  and  $p < .05$  respectively, but not when teacher directiveness was controlled. Perceived teacher directiveness appeared to be somewhat independent of actual teacher directiveness and was correlated with satisfaction ( $r = .34$ ,  $p < .001$ ).*

Researchers have spent considerable time and energy trying to delineate the factors which may enhance student performance and satisfaction in the classroom. One such line of research has been concerned with the relative effects of directive and nondirective teaching styles. In his approach to teaching, the directive teacher tends to be absolute, formal, and structured, while the nondirective teacher tends to be tentative, informal, and

unstructured. Tuckman (1968) views the directive teacher as one for whom extreme concern with order, procedure, and organization affects both the structure of the class and interpersonal relationships with the students. A directive teacher highly structures and plans the course and classroom activity, minimizes informal and group work (and even structures it when it is used), takes absolute responsibility for grades, places a premium on factual knowledge derived from absolute sources, minimizes students' opportunities to learn from their own mistakes, uses absolute and justifiable punishment, and maintains formal teacher-student relationships. A non-directive teacher exhibits the converse of these practices.

Several researchers have sought to demonstrate the superiority of either directive or nondirective teaching to foster greater learning or higher student morale, but the results have been largely inconsistent and inconclusive (Anderson, 1959; McKeachie, 1954; Weiss, Sales, & Bode, 1970). However, as Arlin (1975), Domino (1975), and Gardner (1974) have pointed out, if pupil characteristics are not considered in concert with teacher characteristics it may be falsely concluded that certain teacher variables have little effect upon the teaching outcome under investigation. Student characteristics may mediate the effects of teacher characteristics.

Tuckman (1968) and Weiss et al. (1970) suggested that the effects of teacher directiveness are mediated by student authoritarianism (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950). Tuckman found that nonauthoritarian students expressed greater satisfaction with nondirective teachers while authoritarian students showed little differentiation in their level of satisfaction with one teaching style or the other, in a vocational high school sample. Similarly, nonauthoritarian students earned higher grades from nondirective teachers, but authoritarian students received similar grades under both styles. However, the interaction effects were only found in nonconventional courses and only attained the .10 level of significance.

Weiss et al. used teacher authoritarianism scores as indicators of directiveness to test the same general hypothesis as Tuckman. It is generally assumed that authoritarian teachers use directive teaching styles (McGee, 1955; Maney, 1959). In an academic high school sample, Weiss et al. found that authoritarian students were particularly dissatisfied with nonauthoritarian teachers while satisfaction for nonauthoritarian students was not related to teacher authoritarianism. The findings with grades as the dependent variable were parallel to those with satisfaction, even when the effects of intelligence were taken into account. Again, the interaction effects were only significant at the .10 level.

There is some evidence that dogmatic teachers also tend to be directive (Ager, 1970; Stevenson, 1970; Tosi, Quaranta, & Frumkin, 1968). Dogmatism (Rokeach, 1960) can be regarded as a "general authoritarianism" not oriented to any particular political ideology (Barker, 1963).

The interaction effects of student dogmatism and teacher directiveness apparently have not been researched in any direct way. However, Cohen (1971), in an exploration of tutor-student relationships, found that dogmatic students preferred that the tutor direct studies rather than provide opportunities for the student to direct himself.



Ford (1969) and Appel (1974) postulated that student dogmatism mediates the effects of teacher dogmatism on achievement. In each study, no significant differences were found between the various combinations of teachers and students ranked high and low on the dogmatism dimension.

The Pensacola Z Survey (Jones, 1957) is also purported to measure general authoritarianism. It is comprised of five subscales: rigidity, anxiety, hostility, dependency, and heteronomy. Way (1973) administered the scale to a sample of instructors and students in freshman English. Students, ranked high or low on rigidity, anxiety, and hostility paired with instructors ranked similarly on these subscales, achieved at a higher level than did any group of students with instructors ranked medium. Way pointed out that “although the simple effects reported in the study are not synonymous with interaction effects, they do strongly suggest such interactions” (p. 233).

In summary, the interactive relationships of teacher directiveness and student authoritarianism, and teacher authoritarianism and student authoritarianism, to performance and satisfaction have received only statistically insignificant support; the interactive relationship of teacher directiveness and student dogmatism to either performance or satisfaction has not been specifically researched; the interactive relationship of teacher dogmatism and student dogmatism to performance has been subjected to only quasi interaction analysis, while the relationship to satisfaction has received no attention.

Even if significant interaction effects were reported for any of the four combinations of teacher and student characteristics it would be difficult to assess their validity since teacher directiveness, authoritarianism, and dogmatism appear to be interrelated. Without appropriate controls it might be falsely concluded that teacher personality rather than teaching behavior is the crucial factor, or with controls the “significant” interaction effect might vanish.

The present study was designed to determine which of the relevant interaction effects—teacher directiveness and student authoritarianism, teacher directiveness and student dogmatism, teacher authoritarianism and student authoritarianism, and teacher dogmatism and student dogmatism—account for unique variance in grades and satisfaction.

Intelligence, teacher sex, student sex, and the match of teacher and student sex were considered as covariates in this study since, in addition to the generally accepted relationship between intelligence and achievement, research suggests that teacher ratings are related to intelligence (e.g., Appel, 1974), and achievement and satisfaction are related to teacher sex (e.g., Arnold, 1968; Stasz, Weinberg, & McDonald, 1974), student sex (e.g., Kehle, Bramble, & Mason, 1974; Stasz, Weinberg, & McDonald, 1974), and the match of teacher and student sex (e.g., Costin, Greenough & Menges, 1971; Goebes & Shore, 1975).

### *Method*

#### *Subjects*

Teachers and students in eleventh, twelfth, and thirteenth grade English classes at an Ontario town secondary school participated in this study. The

school has an enrolment of over 1,000 pupils and offers a broad range of academic and vocational courses.

The four male and two female teachers in the study comprised the entire staff teaching English at these grade levels. Two classes of each teacher were selected. Two teachers each taught two eleventh grade classes, two teachers each taught one eleventh grade and one twelfth grade class, one teacher taught two twelfth grade classes, and one teacher taught two thirteenth grade classes.

English was chosen because it can be adequately and comfortably taught using either a directive or nondirective approach, and since the content is more value-laden than a subject such as mathematics, personal values, beliefs, and attitudes would be more apt to be expressed in the course of teaching.

### *Instruments*

Teacher directiveness was measured with the Student Perception of Teacher Style (SPOTS) scale (Tuckman, 1968). Several teacher behaviors are rated by the teacher's students and the mean student score is taken as an index of the teacher's degree of directiveness, lower scores indicating a more directive style. Authoritarianism was measured with a modified version of the F-scale of Adorno et al. (1950) developed by Ezekiel and Athanasiou (Athanasiou, 1968; Ezekiel, 1970) which corrects for response bias, has a test-retest reliability of .86, and avoids the harsh quality of the items of the earlier F-scale (Weiss et al., 1970). Two items which deal with "Americanism" were replaced with two suitable items from another revised F-Scale (Kohn, 1972) because they do not have the same connotations in Canada. Dogmatism was measured with the 40-item Dogmatism Scale (Form E) developed by Rokeach (1960) to tap "general authoritarianism." Adequate reliability and validity for the scale has been demonstrated (e.g., Kerlinger & Rokeach, 1966; Rokeach, 1960; Vacchiano, Strauss, & Hochman, 1969).

As in the Weiss et al. study, two items were employed to measure satisfaction on 6-point Likert scales: (1) I am very satisfied with my present English teacher, and (2) My feelings toward my present English teacher are very favourable.

### *Procedure*

Teachers were asked to complete the F-scale and the Dogmatism Scale after their cooperation had been secured by the head of the English department.

For each class, one period was used solely for the collection of student data. Only the researcher and the students were present. The number of students registered in each class at the beginning of the year had been determined from enrolment records. It was the authors' intention to collect data on about 80% of the students registered in the classes participating in the study. The questionnaires containing the SPOTS scale, the F-Scale, the Dogmatism Scale, and the two satisfaction items were unsystematically distributed and when the quota was reached students without a booklet were excused from the class.



For each student, the following information was determined from school files: student sex, teacher sex, IQ, academic program (four or five year), academic year, English grade on the January report, and the student's English teacher. The anonymity and confidentiality of student and teacher responses and file information was guaranteed through a coding plan.

Results

Of the 196 student questionnaires collected, 22 were discarded because they lacked certian information; compared to the total enrolment, the percentage of usable questionnaires for each of the six teachers' classes ranged from 59% to 83% with the mean being 68%.

Intelligence was measured by the Henmon-Nelson Test of Mental Abilities for 170 students, and the IQ of the four remaining students was measured by the Otis-Beta, the Dominion, the SRA, and the Lorge-Thorndike tests.

Approximately 44.8% of the students were male, the sex of the teacher and the student were matched in 42.5% of the cases, 57.5% of the students were in the eleventh grade, 26.4% were in the twelfth grade, and 16.1% were in the thirteenth grade. Approximately 62.1% of the students were in the five-year program. Descriptive statistics for the major variables are presented in Table 1.

TABLE 1  
DESCRIPTIVE STATISTICS FOR THE SAMPLE

Variable	M	SD	Minimum	Maximum
Grades	64.8	11.7	27	95
Satisfaction	11.3	2.9	2	14
Teacher directiveness <sup>a</sup>	5.12	.49	4.53	5.71
Teacher authoritarianism <sup>a</sup>	82.5	13.1	59	97
Teacher dogmatism <sup>a</sup>	109.3	52.4	108	149
Student authoritarianism	102.6	15.7	68	150
Student dogmatism	163.1	25.9	86	227
Student intelligence	112.1	11.1	86	144

<sup>a</sup> Computed with an N of 6 rather than 174.

A second grades variable, "z-grades," was developed by computing the z-score of each student's grade, using each teacher's students as a subsample. It was thought that z-grades might be a better comparative index of standing across teachers. Satisfaction was uncorrelated with z-grades or grades. As expected, z-grades and grades correlated highly ( $r = .93$ ).

Stepwise multiple regression strategies outlined by Cohen (1968) and Cohen and Cohen (1975) were used to analyze the hypothesized interaction effects. If the product of two independent variables is entered into a regression equation directly after the two independent variables, the significance of the amount of additional variance accounted for in the dependent variable indicates whether or not there is a significant linear interaction between the two independent variables. Accordingly, four

interaction variables were created: teacher directiveness x student authoritarianism, teacher directiveness x student dogmatism, teacher authoritarianism x student authoritarianism, and teacher dogmatism x student dogmatism. Therefore, the initial independent variables were the four product variables, teacher directiveness, teacher authoritarianism, teacher dogmatism, student authoritarianism, student dogmatism, student intelligence, teacher sex, student sex, the match of teacher and student sex, academic program, and two variables carrying the academic year information. The dependent variables were grades, *z*-grades, and satisfaction. All of the variables were transformed to *z*-scores for the regression analysis.

The results of the regression equations reported in Table 2 were deduced in the following manner. Successive predictors were entered into the regression equation in a partially predetermined order. All of the independent variables except those from which the interactions being tested were derived, were considered as covariates and entered first. The independent variables from which the interactions being tested were derived were entered second, and the variables containing the interactions which were being tested were entered last. The amount of variance in the dependent variable accounted for by the total combination of predictors was invariably significant ( $p < .001$ ). The significance of the interaction effects was found by testing the increment in the amount of variance accounted for by the inclusion of the interaction variable or group of interaction variables (Cohen, 1968, p. 435).

In Step 1, the combined interaction effects of teacher directiveness and student authoritarianism, teacher directiveness and student dogmatism, teacher authoritarianism and student authoritarianism, and teacher dogmatism and student dogmatism were found to be insignificant with grades, *z*-grades, or satisfaction as the dependent variable. Therefore, the interactions were not tested individually.

Each student's score on the SPOTS scale was then used to create a "perceived teacher directiveness" variable. It was thought that the failure to find significant interaction effects might have been due to the restricted range of teacher directiveness scores, and also that perceived teacher directiveness might be somewhat independent of actual teacher directiveness. The perceived teacher directiveness variable was correlated with the teacher directiveness variable,  $r(172) = .51, p < .001$ . Employing perceived teacher directiveness rather than teacher directiveness, the correlation with satisfaction rose from .06 to significance,  $r(172) = .34, p < .001$ . In comparison, the correlation with grades dropped from an insignificant -.15 to .02. The corresponding correlation with *z*-grades rose from -.01 to .10. Correlations between perceived teacher directiveness and satisfaction were then computed separately for each teacher. Although numbers ranged from only 18 to 35, three of the correlations were significant and all six were positive:  $r(26) = .67, p < .01$ ;  $r(26) = .54, p < .01$ ;  $r(28) = .40, p < .05$ ;  $r(33) = .24, ns$ ;  $r(16) = .19, ns$ ;  $r(33) = .01, ns$ .

Two product variables were created to carry the interactions of perceived teacher directiveness and student authoritarianism and perceived teacher directiveness and student dogmatism, and were transformed to *z*-scores. The perceived teacher directiveness variable was also standardized. The



TABLE 2  
SUMMARIES OF THE REGRESSION ANALYSES OF THE  
INTERACTION EFFECTS

Step	Interaction(s) Tested	Dependent Variables	R <sup>2</sup> of independent variables entered as covariates	R <sup>2</sup> of covariates plus the independent variables from which the interaction var- iables were derived	R <sup>2</sup> of total number of independent variables	Significance of interaction(s) tested
1	TDIR x SAUTH	Grades	.251	.300	.307	F(4, 157) = .39
	TDIR x SDOG	z-grades	.191	.223	.243	F(3, 158) = 1.37
	TAUTH x SAUTH	Satisfaction	.176	.229	.248	F(4, 157) = .97
	TDOG x SDOG					
2	PTDIR x SAUTH	Grades	.279	.286	.328	F(2, 159) = 4.87**
	PTDIR x SDOG	z-grades	.206	.215	.261	F(2, 159) = 4.95**
		Satisfaction	.188	.314	.360	F(2, 159) = 5.61**
3	PTDIR x SAUTH	Grades	.280	.286	.316	F(1, 160) = 6.94**
		z-grades	.208	.215	.247	F(1, 160) = 6.75*
		Satisfaction	.220	.314	.315	F(1, 160) = .16
	PTDIR x SDOG	Grades	.283	.266	.306	F(1, 160) = 4.50*
		z-grades	.210	.215	.239	F(1, 160) = 4.89*
		Satisfaction	.191	.314	.355	F(1, 160) = 10.07**
4	PTDIR x SAUTH	Grades	.280	.280	.280	Insufficient F-level <sup>a</sup>
		z-grades	.208	.239	.261	F(1, 159) = 4.88*
	PTDIR x SDOG	Grades	.283	.283	.283	Insufficient F-level
		z-grades	.211	.24	.261	F(1, 159) = 3.05
		Satisfaction	.285	.315	.360	F(1, 159) = 11.06**
5a	PTDIR x SAUTH	Grades	.296	.296	.296	Insufficient F-level
	PTDIR x SDOG	z-grades	.219	.229	.268	F(2, 158) = 4.23*
		Satisfaction	.192	.334	.375	F(2, 158) = 5.18**
5b	PTDIR x SAUTH	Grades	.297	.297	.297	Insufficient F-level
		z-grades	.220	.229	.256	F(1, 159) = 5.78*
	PTDIR x SDOG	Grades	.299	.299	.299	Insufficient F-level
		z-grades	.221	.229	.249	F(1, 159) = 4.09*
		Satisfaction	.194	.334	.369	F(1, 159) = 8.66**
5c	PTDIR x SAUTH	z-grades	.222	.249	.268	F(1, 159) = 4.28*
	PTDIR x SDOG	z-grades	.224	.256	.268	F(1, 158) = 2.61
		Satisfaction	.315	.336	.375	F(1, 158) = 9.85**
6	TDIR x SIQ	Grades	.146	.146	.146	Insufficient F-level
	TAUTH x SIQ	z-grades	.074	.223	.230	F(3, 158) = .46
	TDOG x SIQ	Satisfaction	.199	.229	.255	F(3, 158) = 1.80
7a	PTDIR x SIQ	Grades	.170	.286	.289	F(1, 160) = .68
		z-grades	.079	.215	.218	F(1, 160) = .61
		Satisfaction	.206	.314	.333	F(1, 160) = 4.47*
7b	PTDIR x SIQ	Satisfaction	.206	.334	.354	F(1, 159) = 4.87*
7c	PTDIR x SIQ	Satisfaction	.206	.369 <sup>b</sup>	.380	F(1, 158) = .34
7d	PTDIR x SDOG	Satisfaction	.194	.354 <sup>c</sup>	.380	F(1, 158) = 6.66**

Note: TDIR - Teacher directiveness      TAUTH - Teacher authoritarianism  
SAUTH- Student authoritarianism      TDOG - Teacher dogmatism  
SDOG - Student dogmatism              PTDIR - Perceived teacher directiveness  
SIQ - Student intelligence

<sup>a</sup> F-level or tolerance level insufficient for further computation.

<sup>b</sup> The variance accounted for by the interaction of perceived teacher directiveness and student dogmatism is also included.

<sup>c</sup> The variance accounted for by the interaction of perceived teacher directiveness and student intelligence is also included.

\*  $p < .05$

\*\*  $p < .01$

significance of the combined relationship of these two interactions to grades, *z*-grades, and satisfaction was tested in Step 2. The original teacher directiveness variable was not entered into the regression equation. The combined effect was significant ( $p < .01$ ) with each dependent variable.

In Step 3, each interaction effect was tested separately. The interaction of perceived teacher directiveness and student authoritarianism was significantly related to grades,  $F(1, 160) = 6.94$ ,  $p < .01$ , and to *z*-grades. The interaction of perceived teacher directiveness and student dogmatism was significantly related to grades,  $F(1, 160) = 4.50$ ,  $p < .05$ , to *z*-grades, and to satisfaction.

In Step 4, an attempt was made to determine which of the two interactions in Step 3 is the primary interaction in regard to each dependent variable, or if the two interactions have overlapping effects. Each of the significant interaction effects in Step 3 was tested with the other interaction variable entered as a covariate directly before the interaction variable being tested. The interaction of perceived teacher directiveness and student authoritarianism was significantly related to *z*-grades; the interaction of perceived teacher directiveness and student dogmatism was significantly related to satisfaction.

In Step 5, the significant interaction effects found in Step 2, 3, and 4 were tested while the effects of actual teacher directiveness (as determined by Tuckman's recommended usage of the SPOTS scale) were controlled by entering teacher directiveness as a covariate. Tested as a pair in Step 5a, the interactions were significantly related to *z*-grades and to satisfaction. Tested separately in Step 5b, the interaction of perceived teacher directiveness and student authoritarianism was significantly related to *z*-grades, and the interaction of perceived teacher directiveness and student dogmatism was significantly related to *z*-grades and satisfaction. In Step 5c, with the effects of the interaction of perceived teacher directiveness and student dogmatism controlled, the interaction of perceived teacher directiveness and student authoritarianism was significantly related to *z*-grades,  $F(1, 159) = 4.28$ ,  $p < .05$ . With the effects of the interaction of perceived teacher directiveness and student authoritarianism controlled, the interaction of perceived teacher directiveness and student dogmatism was significantly related to satisfaction,  $F(1, 158) = 9.85$ ,  $p < .01$ .

Since intelligence has been found to correlate with authoritarianism (e.g., Weiss et al., 1970) and dogmatism and academic aptitude have been found to correlate (e.g., Steininger, 1970), an attempt was made to determine whether or not the present significant interaction effects could be explained by the interaction of perceived teacher directiveness and student intelligence. Four product variables were created—teacher directiveness  $\times$  student intelligence, teacher authoritarianism  $\times$  student intelligence, teacher dogmatism  $\times$  student intelligence, and perceived teacher directiveness  $\times$  student intelligence—and transformed to *z*-scores.

It was first determined in Step 6 that the combined effect of the interactions of teacher directiveness and student intelligence, teacher authoritarianism and student intelligence, and teacher dogmatism and student intelligence was insignificant in relation to grades, *z*-grades, and satisfaction. In Step 7a, it was found that the interaction of perceived



teacher directiveness and student intelligence was significantly related to satisfaction, but not to grades or *z*-grades, and remained so when teacher directiveness was entered as a covariate in Step 7b. In Step 7c, the perceived teacher directiveness  $\times$  student dogmatism variable was entered directly before the perceived teacher directiveness  $\times$  student intelligence variable, and it was found that the interactive relationship of perceived teacher directiveness and student intelligence to satisfaction was rendered insignificant. However, in Step 7d, when the perceived teacher directiveness  $\times$  student intelligence variable was entered directly before the perceived teacher directiveness  $\times$  student dogmatism variable, the interactive relationship of perceived teacher directiveness and student dogmatism to satisfaction remained significant.

In summary, with grades as the dependent variable in Step 3, there was a significant interaction between both perceived teacher directiveness and student authoritarianism and perceived teacher directiveness and student dogmatism. Neither of these interaction effects was significant in Step 4 when each of the interaction variables was tested with the other interaction variable entered before it in the regression equation and consequently controlled. Therefore, the two significant interaction effects of Step 3 are presented in Figure 1<sup>1</sup> and Figure 2.

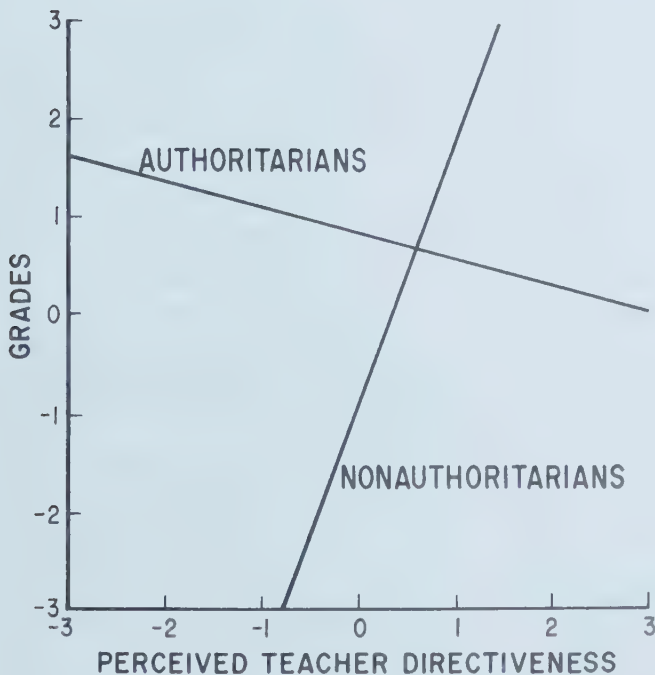


FIGURE 1

The interactive relationship of perceived teacher directiveness and student authoritarianism to grades in Step 3.

*Note.* The standardized form of the regression equation is:  $z = .86x + 1.23y - 1.50xy$ , where  $z$  = Grades,  $x$  = Student authoritarianism, and  $y$  = Perceived teacher directiveness.

The interactive relationship of perceived teacher directiveness and student authoritarianism to *z*-grades maintained even when teacher directiveness was controlled in Step 5c, and is presented in Figure 3. Similarly, the interactive relationship of perceived teacher directiveness and student dogmatism to satisfaction was also significant in Step 5c and is presented in Figure 4.

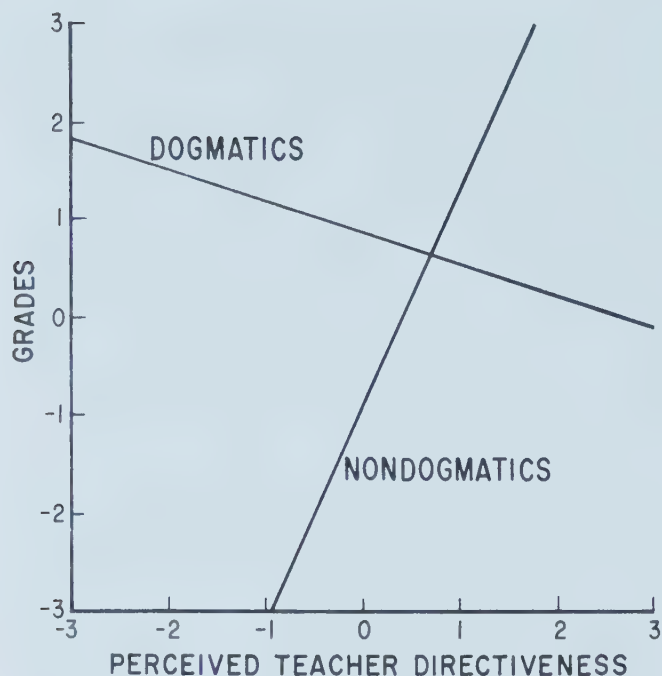


FIGURE 2

The interactive relationship of perceived teacher directiveness and student dogmatism to grades in Step 3. *Note.* The standardized form of the regression equation is:  $z = .86x + .95y - 1.27xy$ , where  $z$  = Grades,  $x$  = Student dogmatism, and  $y$  = Perceived teacher directiveness.

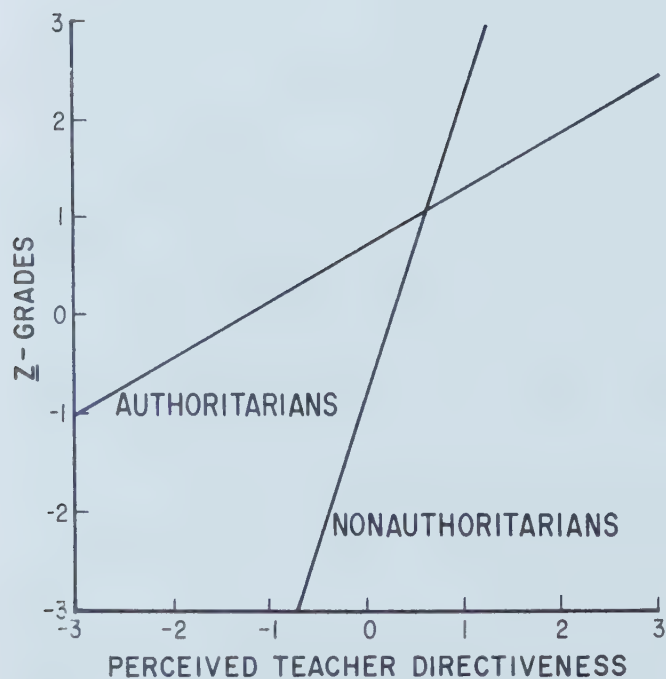


FIGURE 3

The interactive relationship of perceived teacher directiveness and student authoritarianism to z-grades in Step 5c. *Note.* The standardized form of the regression equation is:  $z = 1.83x + .73y - 1.26xy$ , where  $z$  = z-grades,  $x$  = Perceived teacher directiveness, and  $y$  = Student authoritarianism.



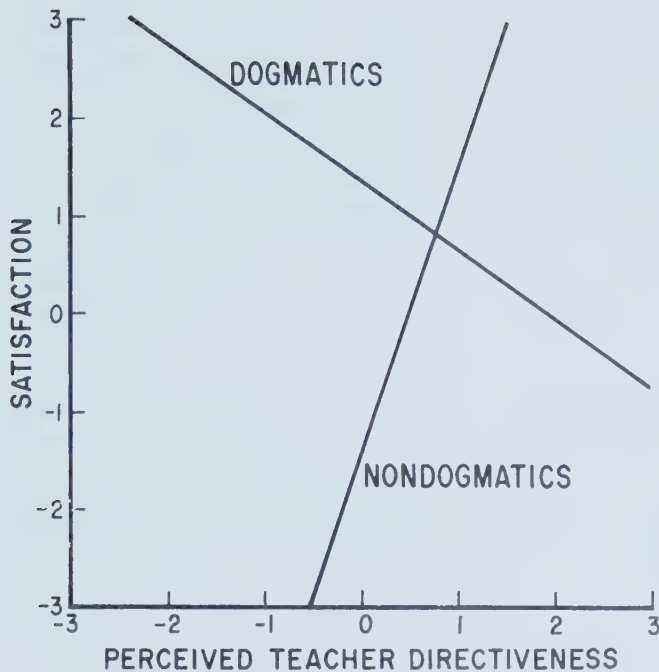


FIGURE 4

The interactive relationship of perceived teacher directiveness and student dogmatism to satisfaction in Step 5c.

*Note.* The standardized form of the regression equation is:  $z = 1.33x + 1.14y - 1.84xy$ , where  $z$  = Satisfaction,  $x$  = Student dogmatism, and  $y$  = Perceived teacher directiveness.

### Discussion

The expected interactive relationships of teacher directiveness and student authoritarianism, teacher directiveness and student dogmatism, teacher authoritarianism and student authoritarianism, and teacher dogmatism and student dogmatism to grades,  $z$ -grades, and satisfaction were insignificant. However, several interactions involving teacher directiveness were significant when perceived teacher directiveness replaced the conventional measure of teacher directiveness.

Nonauthoritarian or nondogmatic students earned higher grades when they perceived the teacher to be nondirective; authoritarian or dogmatic students earned higher grades when they perceived the teacher to be directive. Both interactions accounted for much of the same variance in grades, since each was insignificant when the other was statistically controlled. Each interaction effect was also apparently contingent upon the actual level of teacher directiveness, since each was insignificant when teacher directiveness was statistically held constant.

Nonauthoritarian and authoritarian students both earned higher  $z$ -grades when they perceived the teacher to be nondirective, but the effect was more pronounced for nonauthoritarian students. This interaction effect remained significant when teacher directiveness was controlled.

The more valid dependent variable to represent student performance appears to be  $z$ -grades rather than grades. Grades can vary across teachers because of each teacher's accustomed mean, range, and distributional pattern. This variability, which may be unrelated to actual performance, is corrected with  $z$ -grades. The standardized version can be used as a comparative index of relative standing in the classes of each teacher and across the classes of all teachers.

Nondogmatic students were more satisfied when they perceived the teacher to be nondirective, while dogmatic students were more satisfied

when they perceived the teacher to be directive. This interactive relationship remained significant when teacher directiveness was controlled.

The preceding four interaction effects could not be accounted for by perceived teacher directiveness, teacher authoritarianism, teacher dogmatism, student authoritarianism, student dogmatism, student intelligence, teacher sex, student sex, the match of teacher and student sex, academic year, academic program, the interaction of perceived teacher directiveness and student intelligence, or the additive effects of these variables.

The step-wise regression procedure used here provides a conservative test of the effects of the interactions on grades, z-grades, and satisfaction. Only after the variance accounted for by the variables in the foregoing paragraph was removed from the dependent variable, were the interaction terms entered into the regression equation. These interaction terms had to account for significant amounts of additional criterion variance to reach statistical significance. It was not sufficient that they merely be related to the same variance accounted for by a prior term. Hence, the amount of unique variance accounted for by the interaction terms may seem somewhat small.

Teacher authoritarianism and dogmatism scores were low and had a narrow range. As Tuckman (1968) and Weiss et al. (1970) similarly argued, these factors may have attenuated interaction effects involving teacher authoritarianism and dogmatism. The restricted range of teacher directiveness scores may have also attenuated interaction effects involving teacher directiveness. However, it is the present authors' position that actual teacher directiveness is important in the present context only to the extent that it influences each student to perceive the teacher as being directive or nondirective. It might also be speculated that teacher authoritarianism and dogmatism are relevant in the present context to the extent that they influence perceived teacher authoritarianism and dogmatism.

There is evidence that perceived teacher directiveness is somewhat independent of actual teacher directiveness. There was a wide range of perceived teacher directiveness scores in the classes of each teacher, and for the most part, these scores were systematically related to satisfaction with each teacher. Students expressed greater satisfaction when they perceived the teacher to be nondirective. Since actual teacher directiveness (and not simply that estimated by the mean SPOTS scale score) was constant for each teacher, perceived teacher directiveness must be related to satisfaction in its own right. Further similar evidence is provided by the significant interactive relationships involving perceived teacher directiveness when the effects of teacher directiveness (as estimated by the mean SPOTS scale score) were statistically controlled.

The present study has helped to clarify the interactive relationship of teacher directiveness and student authoritarian tendencies to achievement and morale. Hitherto, significant interaction effects have not been reported. The expected nature of such interaction effects seems to have been generally correct, but one must consider that the teaching outcome is at least somewhat dependent on the students' perceptions of the classroom situation. Apparently, student authoritarianism interacts with perceived teacher directiveness in the determination of grades, while student dogmatism



interacts with perceived teacher directiveness in the determination of satisfaction.

The authors wish to thank Mr. G. Allan Snider, Director of Education; Mr. E. H. B. "Burt" Garrett, Principal; Mr. John S. Morgan, Vice-Principal; Mr. Haddon Meldrum and Mrs. Glenna Moore of the Guidance Department; and Mr. Peter J. Maule of the English Department at Parry Sound High School for their cooperation and assistance in the collection of the data.

Requests for reprints should be sent to Ronald J. Fisher, Department of Psychology, University of Saskatchewan, Saskatoon, Saskatchewan, S7N 0W0.

<sup>1</sup> In Figure 1, 2, 3, and 4, "authoritarians" and "dogmatics" are 1 *SD* above the *M*, while "nonauthoritarians" and "nondogmatics" are 1 *SD* below the *M*. The *M* of 0 was substituted for each of the variables in the original regression equation which are not present in each figure.

### References

- Adorno, T., Frenkel-Brunswik, E., Levinson, D., & Sanford, N. *The authoritarian personality*. New York: Harper, 1950.
- Ager, M. Dogmatism and the verbal behavior of student teachers. *Journal of Teacher Education*, 1970, 21, 179-183.
- Anderson, R. C. Learning in discussions: A resume of the authoritarian-democratic studies. *Harvard Educational Review*, 1959, 29, 201-215.
- Appel, D. A study of the relationships of student ratings of teachers, achievement gains and the dogmatism of teachers and students. *Dissertation Abstracts International*, 1974, 34(8-A), 4853.
- Arlin, M. The interaction of locus control, classroom structure, and pupil satisfaction. *Psychology in the Schools*, 1975, 12, 279-286.
- Arnold, R. D. The achievement of boys and girls taught by men and women teachers. *Elementary School Journal*, 1968, 68, 367-372.
- Athanasiou, R. Technique without mystique: A study of authoritarianism in engineering students. *Educational and Psychological Measurement*, 1968, 28, 1181-1188.
- Barker, E. N. Authoritarianism of the political right, center, and left. *Journal of Social Issues*, 1963, 19, 63-74.
- Cohen, J. Multiple regression as a general data-analytic system. *Psychological Bulletin*, 1968, 70, 426-443.
- Cohen, J., & Cohen, P. *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, N.J.: Lawrence Erlbaum, 1975.
- Cohen, U. Personality, school experience and preferences for tutor-student relationships in colleges of education. *Human Relations*, 1971, 24, 349-358.
- Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research*, 1971, 41, 511-535.
- Domino, G. Let the punishment fit the crime: Teacher-student interactions. *Journal of Educational Research*, 1975, 68, 8-11.
- Ezekiel, R. S. Authoritarianism, acquiescence and field behavior. *Journal of Personality*, 1970, 38, 31-42.
- Ford, J. A study of the relationship between belief systems and cognitive functioning in children of grades four, five, and six. *Dissertation Abstracts*, 1969, 30(1-A), 187-188.
- Gardner, P. L. Research on teacher effects: Critique of a traditional paradigm. *British Journal of Educational Psychology*, 1974, 44, 123-130.
- Goebes, D. D., & Shore, M. F. Behavioral expectations of students as related to the sex of the teacher. *Psychology in the Schools*, 1975, 12, 222-224.
- Jones, M.B. The Pensacola Z Survey: A study in the measurement of authoritarian tendency. *Psychological Monographs*, 1957, 71, (Whole No. 45).
- Kehle, T. J., Bramble, E. J., & Mason, E. J. Teacher's expectation: Ratings of student performance as biased by student characteristics. *Journal of Experimental Education*, 1974, 43, 54-60.

- Kerlinger, F., & Rokeach, M. The factorial nature of the F and D Scales. *Journal of Personality and Social Psychology*, 1966, 4, 391-399.
- Kohn, P. M. The Authoritarianism-Rebellion Scale: A balanced F-Scale with left-wing reversals. *Sociometry*, 1972, 35, 176-189.
- Maney, A. C. The authoritarianism dimension in student evaluations of faculty. *Journal of Educational Sociology*, 1959, 32, 226-231.
- McGee, H. M. Measurement of authoritarianism and its relation to teachers' classroom behavior. *Genetic Psychology Monographs*, 1955, 52, 89-146.
- McKeachie, W. J. Student centered versus teacher centered instruction. *Journal of Educational Psychology*, 1954, 45, 143-150.
- Rokeach, M. *The open and closed mind*. New York: Basic Books, 1960.
- Stasz, C., Weinberg, S., & McDonald, F. J. The influence of sex of student and sex of teacher on student's achievement and evaluation of the teacher. *Catalog of Selected Documents in Psychology*, 1974, 54. (Ms. No. 421)
- Steininger, M. Aptitude, dogmatism, and college press as codeterminants of academic achievement. *Journal of Social Psychology*, 1970, 80, 229-230.
- Stevenson, W. A. Authoritarian and equalitarian attitudes of classroom teachers. *Dissertation Abstracts International*, 1970, 31(4-A), 1556.
- Tosi, D., Quaranta, J. J., & Frumkin, R. M. Dogmatism and student teacher perceptions of ideal classroom leadership. *Perceptual and Motor Skills*, 1968, 27, 750.
- Tuckman, B. W. A study of the effectiveness of directive versus nondirective vocational teachers as a function of student characteristics and course format. *United States Office of Education Research Project No. 6-2300*, Rutgers University, 1968.
- Vacchiano, R.B., Strauss, P. S., & Hochman, L. The open and closed mind: A review of dogmatism. *Psychological Bulletin*, 1969, 71, 261-273.
- Way, J. G. The relation of student and teacher traits of a authoritarianism to student achievement in English. *Journal of Psychology*, 1973, 85, 229-234.
- Weiss, R. L., Sales, S. M., & Bode, S. Student authoritarianism and teacher authoritarianism as factors in the determination of student performance and attitudes. *Journal of Experimental Education*, 1970, 38, 83-87.



M. HAUGHEY  
E. A. HOLDAWAY  
J. M. SMALL

*The University of Alberta*

## Consultative Practices of Elementary School Teachers: Some Research Findings

*This study sought detailed information about the consultative needs and practices of elementary school teachers. Questionnaires and interviews were used with 80 teachers in three elementary schools in one school system in Alberta. Assistance was most commonly desired in tasks related to special needs of students (e.g., diagnosis of learning difficulties), selecting instructional materials, operating audio-visual equipment, and developing course outlines. An extensive pattern of consultation among teachers was revealed, and teachers were generally satisfied with the assistance that they received. However, teachers with more years of post-secondary education or with more years of experience were less satisfied with this assistance.*

*Several recommendations are made. Principals should assess the consultative needs and capabilities of their staff more carefully, and school systems should probably emphasize specialist consultation services on the system level rather than the general consultation which can usually be adequately provided within schools. (Dr. Haughey is a sessional lecturer, Dr. Holdaway is a Professor, and Dr. Small is Associate Professor in the Department of Educational Administration, The University of Alberta.)*

Teachers, in common with people in most occupational groups, frequently seek opportunities to discuss aspects of their work with others, but the means by which such discussions occur vary widely. Three traditional models of teacher consultation may be referred to as "big brother," "fire fighter," and "casual visitor." In all cases, a superordinate-subordinate relationship is assumed, with the teacher, inevitably, assuming the latter role, while the superordinate is either the principal or someone designated by the system as consultant or supervisor.

The "big brother" approach is most often seen in relation to an inexperienced or new member on the school staff. The principal or other

supervisor is seldom far from the classroom and is ready to offer advice on a multitude of matters, whether it is solicited or not. By contrast, the "fire fighter" is called in, or rushes in, when emergency situations arise which the teacher is incapable of handling, such as disruptive pupils or untenable working situations. The "casual visitor" drops in from time to time as a means of informal quality control, to maintain communications, and occasionally to inform the teacher of prospects and programs which may be of interest.

These models of supervision all rest upon the inherent assumptions which may limit their potential usefulness to a teaching force which is becoming increasingly professionalized. The assumptions are that teachers as a body do not recognize their own needs and lack the motivation to initiate consultation, and as a corollary, that teachers are not viewed as having resources which may be valuable to their colleagues.

Many commentators have questioned the assumptions behind hierarchical supervision, and have looked towards a more open mode of communication among educators at all levels (for example: Lucio, 1967; Sergiovanni & Starratt, 1971; MacKay, 1971; and Miklos, 1971).

More and more, teachers are perceived to be the best judges of what they need to know, how and from whom this knowledge should be acquired, and from the information so obtained, what can best be integrated with their own teaching styles. The study reported here provides a further test of the validity of this perception by examining the dynamics of consultative encounters between in-school personnel at the elementary level in a small urban school system in Alberta.

### *Identification of Consultants*

A number of studies have focused on the personnel with whom teachers discuss their concerns and those perceived as valuable sources of assistance. From a general review of studies in these areas, Holdaway (1971) noted that (1) only a small percentage of a school system's staff are designated as consultants, (2) these people are unable to meet all the demands for their services, (3) principals have not enough expertise to provide the kind of specialist help required by teachers, (4) principals are key figures in the provision of a supportive, facilitating and motivating climate, and (5) the major task of the consultant should be to help teachers solve problems which the teachers have identified.

Parsons (1971) asked 697 elementary teachers in West Central Ontario to identify and to rate the effectiveness of supervisory personnel who affected their classroom instruction. Principals were rated as being the most influential and most effective; they were followed in order by other teachers, resource teachers, and program consultants. However in three separate studies (McGillivray, 1966; Moller, 1968; and Milne, 1968), beginning teachers reported that they received most help from fellow teachers and department heads, and least help from principals and central office personnel. The importance of other teachers as sources of consultation is apparent from these studies, a conclusion which led Parsons (1971) to suggest that:

... greater opportunities be provided for teacher interaction in school systems by



restructuring teacher roles. Teachers should be given greater freedom from their "in-class" responsibilities to share new ideas and techniques with their colleagues. (p. 8)

### *Identification of Needs*

Research in identification of the consultative needs of teachers has been almost completely confined to first-year teachers. Moller (1968) and Franc (1970) found that a variety of instructional problems were encountered by beginning teachers, and that these were related to concerns of personal competence. Problems occurred in student motivation, grading, meeting individual differences, discipline, working with low-ability students, finding a satisfactory speed for covering course materials, making satisfactory tests, lack of teaching creativity, and a tendency to "talk over students' heads."

Beyond the first years, Fuller and Parsons (1968) and Fuller (1969) suggested that different teacher concerns emerge after different periods of experience. Kass (1975) further developed this idea and postulated a three-stage concern sequence for practising teachers, with self-adequacy being predominant in Stage 1, content and methodology in Stage 2, and concerns with individual student needs in Stage 3.

In general the literature supports the ideal of continuous development throughout the teacher's career in a work environment characterized by the free flow of information between colleagues.

### *Method*

Haughey (1976) sought to describe consultative practices in selected Alberta schools. In particular, teachers were asked to provide information concerning their consultative needs, the personnel whom they consulted, their perceptions of themselves as consultants, their satisfaction with the consultative assistance which they received, and any reasons for not seeking assistance.

### *Population*

The population for the study consisted of all 80 teachers in three elementary schools in one school system in Alberta. This particular system was chosen because some preliminary research had already been conducted in the system, because interviews could be conducted in the time available with all 80 teachers, and because the superintendent and staff were interested and cooperative. This population was not considered to be a sample for inferential purposes of either all elementary teachers in past or future years in the system, or elementary teachers in other school systems. An in-depth approach was chosen in order to obtain detailed information about all teachers in one system: this could lead to hypotheses which may be tested later using comprehensive sampling procedures.

### *Instrument*

Data were collected by means of (1) a questionnaire and (2) an interview schedule designed to clarify and elaborate upon written responses in the questionnaire. Twenty-one task areas were arbitrarily grouped into four categories called "curriculum and program," "instructional," "audio-visual technology," and "special student needs."

The *questionnaire* was based on one used by Plamondon (1973) in earlier research on consultation in the same school system. For each task area, teachers were asked to check if they had desired consultative assistance and whether this had occurred more than once during the school year. Some demographic data were also requested. To ensure that the questions were valid, clear and unambiguous, the questionnaire was distributed to five practicing teachers, three university professors, five graduate students, and to the superintendent, assistant superintendent and three principals in the participating system. Their reactions led to certain changes being made in content and wording: the amended questionnaire was then completed by four elementary teachers in a non-participating elementary-junior high school in the same system.

During the *interviews*, teachers were asked to state their particular concerns, their times of occurrence, the people consulted, and ratings of assistance received from each consultant. They were also asked to identify the staff members for whom they had provided assistance, the task areas in which they had provided assistance, and the frequency of their provision of assistance. The interview schedule was tested and amended using the same procedure as outlined above for the questionnaire.

### *Data analysis*

The raw data consisted of frequency counts, most of which were converted into percentage frequency distributions. Because a population of teachers was used rather than a sample, inferential statistics were not appropriate. Rather, comparisons were made between the size of selected percentage frequencies, with arbitrary decisions being made about the substantial significance of differences in percentages. This method did not allow for projections to be made beyond the population being studied.

## *Results*

### *Areas in Which Teachers Desired Consultative Assistance*

Over 50 percent of the teachers had expressed a desire for consultative assistance in each of seven task areas (Table 1). Fifty-five percent desired assistance in "developing course outlines," 73 percent in "selecting instructional materials," and 64 percent in "operating audio-visual equipment." In all four task areas listed in the special student needs section, over 50 percent of teachers desired assistance. Eighty-five percent of teachers mentioned "diagnosis of learning difficulties," which was the highest percentage of teachers desiring assistance in any task area: the other special student needs task area percentages were "developing remedial programs" (75 percent), "obtaining student-background information" (54 percent), and "solving teacher-pupil problems" (51 percent).

In all, 689 concerns were mentioned by teachers; the highest percentage of these (10 percent) was in the area of "diagnosis of learning difficulties." Nine percent of concerns were in "developing remedial programs," eight percent in "selecting instructional materials," seven percent in "operating audio-visual equipment," and six percent in each of "developing course outlines," "obtaining student-background information," and "solving teacher-pupil problems."



TABLE 1  
PERCENTAGE OF TEACHERS DESIRING CONSULTATIVE ASSISTANCE IN  
EACH TASK AREA AND PERCENTAGE DISTRIBUTION OF CONCERNS  
IN EACH TASK AREA FOR ALL TEACHERS

TASK AREA Description	Teachers (N=80) %	Concerns (N=689) %
CURRICULUM AND PROGRAM		
Developing course outlines	55.0 <sup>a</sup>	6.4 <sup>b</sup>
Selecting instructional materials	72.5	8.4
Developing unit/lesson plans	47.5	5.5
Developing instructional materials	47.5	5.5
Evaluating lessons/units	27.5	3.2
Interpreting curriculum guides	16.3	1.9
INSTRUCTIONAL		
Developing questioning techniques	13.8	1.6
Teaching concepts, constructs, generalizations	35.0	4.1
Planning role-playing	15.0	1.7
Developing/using small group activities	35.0	4.1
Planning individualized instruction	46.3	5.4
Using problem-solving/inquiry/discovery method	13.8	1.6
Leading discussions	11.3	1.3
Grouping for instruction	35.0	4.1
AUDIO-VISUAL TECHNOLOGY		
Operating audio-visual equipment	63.8	7.4
Using kits, charts, games, simulations	35.0	4.1
Developing own audio-visual materials	22.5	2.6
SPECIAL STUDENT NEEDS		
Diagnosis of learning difficulties	85.0	9.9
Developing remedial programs	75.0	8.9
Obtaining student-background information	53.8	6.3
Solving teacher-pupil problems	51.3	6.0

<sup>a</sup> 55% of the 80 teachers in the study desired consultative assistance in Task Area "Developing course outlines."  
<sup>b</sup> Of the 689 concerns mentioned by the 80 teachers in the study, 6.4% were in Task Area "Developing course outlines."

*Education and experience.* The responses indicating teacher concerns in each task area were examined on the basis of years of post-secondary education (Table 2). Greater percentages of teachers with three or more years of post-secondary education desired assistance on more task areas than did teachers with one or two years of post-secondary education. In contrast, as shown in Table 3, lower percentages of teachers with over ten years of teaching experience desired consultative assistance than did teachers with ten or fewer years of teaching experience.

*Recurring concerns.* Teachers were asked to specify for each concern the time of year it occurred. Where the concern emerged more than once it was labelled a recurring concern. The majority of concerns in all areas except "interpreting curriculum guides" were recurring concerns.

*Frequency of consultation.* The number of times each teacher was consulted in each task area and in each section was also examined (Table 4).

TABLE 2  
PERCENTAGE OF TEACHERS DESIRING CONSULTATIVE ASSISTANCE IN  
EACH TASK AREA, BY YEARS OF POST-SECONDARY EDUCATION

TASK AREA Description	Years of Post-Secondary Education			
	1-2 %	3 %	4 %	5-6 %
CURRICULUM AND PROGRAM				
Developing course outlines	30.0	58.8	59.0	57.1
Selecting instructional materials	70.0	82.4	62.2	71.4
Developing unit/lesson plans	40.0	58.8	53.9	21.4
Developing instructional materials	10.0	76.5	48.7	35.7
Evaluating lessons/units	10.0	41.2	25.6	28.6
Interpreting curriculum guides	10.0	11.8	20.5	14.3
INSTRUCTIONAL				
Developing questioning techniques	10.0	5.9	15.4	21.4
Teaching concepts, constructs, generalizations	30.0	47.1	33.3	28.6
Planning role-playing	10.0	29.4	12.8	7.1
Developing/using small group activities	20.0	52.9	38.5	14.3
Planning individualized instruction	20.0	76.5	46.2	28.6
Using problem-solving/inquiry/discovery method	0.0	23.5	12.8	14.3
Leading discussions	0.0	17.7	12.8	7.1
Grouping for instruction	20.0	35.3	38.5	35.7
AUDIO-VISUAL TECHNOLOGY				
Operating audio-visual equipment	70.0	70.6	59.0	64.3
Using kits, charts, games, simulations	10.0	47.1	38.5	28.6
Developing own audio-visual materials	10.0	35.3	15.4	35.7
SPECIAL STUDENT NEEDS				
Diagnosis of learning difficulties	90.0	88.2	84.6	78.6
Developing remedial programs	70.0	94.1	76.9	50.0
Obtaining student-background information	40.0	76.5	46.2	57.1
Solving teacher-pupil problems	50.0	52.9	56.4	35.7
N				
	10	17	39	14
	100.0	100.0	100.0	100.0

In one school the principal, assistant principal, guidance counsellor, two teachers who were involved in remedial work, a Grade 5 teacher and a Grade 6 teacher were each consulted on at least 30 occasions between September and February. In the second school, the personnel consulted most often were the principal, the guidance counsellor, the Grade 1 teacher aide (a fully qualified teacher), two Grade 1 teachers and the remedial-reading teacher. In the third school the assistant principal was consulted on 64 occasions, while other personnel consulted at least 23 times were the principal, the remedial-reading teacher and a Grade 3 teacher who provided most assistance with audio-visual problems.

*Rating of Consultative Assistance*

Teachers generally were satisfied with the information which they received from the teachers who provided assistance for them. Of the 80 teachers, only five had concerns for which they were unable to find a satisfactory answer. One teacher was new to the staff and to the system in



TABLE 3  
PERCENTAGE OF TEACHERS DESIRING CONSULTATIVE ASSISTANCE IN  
EACH TASK AREA, BY TOTAL YEARS OF TEACHING EXPERIENCE

Description	Total Years of Teaching Experience				
	1-3	4-6	7-10	11-20	Over 20
	%	%	%	%	%
CURRICULUM AND PROGRAM					
Developing course outlines	68.0	72.2	54.6	31.6	28.6
Selecting instructional materials	80.0	77.8	90.9	52.6	57.1
Developing unit/lesson plans	72.0	44.4	63.6	21.1	14.3
Developing instructional materials	72.0	44.4	45.5	31.6	14.3
Evaluating lessons/units	36.0	33.3	27.3	15.8	14.3
Interpreting curriculum guides	28.0	11.1	27.3	5.3	0.0
INSTRUCTIONAL					
Developing questioning techniques	20.0	22.2	18.2	0.0	0.0
Teaching concepts, constructs, generalizations	44.0	44.4	18.2	36.8	0.0
Planning role-playing	20.0	22.2	18.2	5.3	0.0
Developing/using small group activities	44.0	33.3	54.6	26.3	0.0
Planning individualized instruction	44.0	50.0	72.7	47.4	0.0
Using problem-solving/inquiry/discovery method	8.0	27.8	9.1	10.5	14.3
Leading discussions	28.0	5.6	9.1	0.0	0.0
Grouping for instruction	40.0	44.4	45.5	26.3	0.0
AUDIO-VISUAL TECHNOLOGY					
Operating audio-visual equipment	84.0	66.7	63.6	42.1	42.9
Using kits, charts, games, simulations	56.0	38.9	36.4	15.8	0.0
Developing own audio-visual materials	36.0	22.2	27.3	0.0	28.6
SPECIAL STUDENT NEEDS					
Diagnosis of learning difficulties	84.0	77.8	100.0	89.5	71.4
Developing remedial programs	84.0	61.1	90.9	84.2	28.6
Obtaining student-background information	64.0	55.6	72.7	36.8	28.6
Solving teacher-pupil problems	52.0	44.4	81.8	42.1	42.9
N	25	18	11	19	7
	100.0	100.0	100.0	100.0	100.0

January, one teacher was unhappy at the level of interaction with her fellow grade teacher, and three were specialists with no other teachers teaching that subject in the school.

In only 123 of the 1545 consultations were the consultants unable to provide information that was satisfactory for the clients. The distribution of teachers' ratings of consultative assistance in each task area is given in Table 5.

In general, teachers rated the consultative assistance which they obtained as "satisfactory" or "very satisfactory." The area in which they were least satisfied was "interpreting curriculum guides" where one-third of the ratings were "not very satisfactory." Teachers also had difficulty finding satisfactory information about "using problem solving/inquiry/discovery method," with 25 percent of ratings being "not very satisfactory."

Teachers' satisfaction tended to decrease with increasing years of post-secondary education. Teachers with one or two years of post-secondary

TABLE 4  
NUMBER OF TIMES EACH TEACHER WAS NAMED AS CONSULTANT

Teacher No.	No. of Times	Teacher No.	No. of Times	Teacher No.	No. of Times
<u>School 1</u>					
4	35	16	15	28	35
5	4	17	13	29	10
6	19	18	5	30	17
7	12	19	27	31	11
8	49	20	14	32	58
9	6	21	12	33	64
10	3	22	5	34	7
11	8	23	17	35	11
12	12	24	26	36	2
13	17	25	32	37	12
14	30	26	20	38	10
15	13	27	30		
<u>School 2</u>					
4	57	12	8	20	0
5	23	13	8	21	5
6	8	14	10	22	1
7	23	15	14	23	24
8	13	16	9	24	4
9	7	17	0	25	2
10	6	18	5	26	15
11	5	19	25	27	25
<u>School 3</u>					
4	17	12	12	20	23
5	2	13	10	21	9
6	15	14	17	22	17
7	11	15	7	23	29
8	14	16	11	24	10
9	64	17	1	25	12
10	17	18	42		
11	4	19	6		

Note. Mean number of times teachers were named as consultants: School 1 = 16.5  
School 2 = 12.6  
School 3 = 15.9

education rated 58 percent of their consultations as “very satisfactory”; teachers with five or six years of post-secondary education indicated that only 33 percent of their consultations were “very satisfactory.” When responses were classified according to years of teacher experience, the pattern was similar, with the percentage of consultations in the “not very satisfactory” category rising from 8 percent for teachers with one to three years of experience to 26 percent for teachers who had over 20 years of teaching experience.

Dissatisfaction with peer consultation appeared, then, to be related to years of post-secondary education and years of experience. A similar kind of relationship existed in regard to the expressed need for consultative help. Ideally teachers expressing a high need for consultation would also indicate a high level of satisfaction, but as Figure 1 shows, this was true only for teachers iwth less experience when their level of formal education is disregarded. Teachers with more post-secondary education, while ex-



TABLE 5  
DISTRIBUTION OF TEACHERS' RATINGS OF CONSULTATIVE ASSISTANCE IN EACH TASK AREA

TASK AREA Description	Rating of Assistance			No. of ratings
	Very satis- factory %	Satis- factory %	Not very satis- factory %	
CURRICULUM AND PROGRAM				
Developing course outlines	40.3 <sup>a</sup>	52.2	7.5	134
Selecting instructional materials	56.2	40.9	2.9	137
Developing unit/lesson plans	34.4	62.2	3.3	90
Developing instructional materials	44.9	49.4	5.6	89
Evaluating lessons/units	30.0	66.7	3.3	30
Interpreting curriculum guides	11.1	55.6	33.3	27
INSTRUCTIONAL				
Developing questioning techniques	22.2	66.7	11.1	9
Teaching concepts, constructs, generalizations	23.0	65.6	11.5	61
Planning role-playing	18.2	81.8	0.0	11
Developing/using small group activities	42.1	45.6	12.3	57
Planning individualized instruction	40.3	44.2	15.6	77
Using problem-solving/inquiry/discovery method	60.0	15.0	25.0	20
Leading discussions	100.0	0.0	0.0	4
Grouping for instruction	36.8	52.6	10.5	76
AUDIO-VISUAL TECHNOLOGY				
Operating audio-visual equipment	69.3	30.7	0.0	75
Using kits, charts, games, simulations	58.1	41.9	0.0	62
Developing own audio-visual materials	44.8	51.7	3.5	29
SPECIAL STUDENT NEEDS				
Diagnosis of learning difficulties	43.9	44.4	11.7	214
Developing remedial programs	44.2	45.7	10.1	129
Obtaining student-background information	50.9	46.4	2.7	112
Solving teacher-pupil problems	40.2	51.0	8.8	102

<sup>a</sup> 40.3% of the ratings in "Developing course outlines" were in the "Very Satisfactory" category.

FIGURE 1  
POSSIBLE RELATIONSHIPS BETWEEN NEED FOR AND SATISFACTION WITH CONSULTATIVE PROCEDURES FOR TEACHERS CLASSIFIED BY (1) MORE OR LESS EXPERIENCE AND (2) MORE OR LESS POST-SECONDARY EDUCATION.

		SATISFACTION	
		High	Low
N E E D	High	less experience	more post-secondary education
	Low	less post-secondary education	more experience

pressing a high need for consultation, regarded peer consultation as less satisfactory than did less qualified teachers.

In all, 27 teachers had concerns for which they did not seek assistance. They mentioned thirteen reasons for not doing so. These reasons fell into three main categories: (1) those where the concern was not crucial but was one where the teacher would like more information; (2) those where the teacher either had no time, no materials, no one to ask or did not know whom to ask; and (3) those where the information became available without seeking assistance. Reasons listing the concern as not crucial were given 24 times, a lack of knowledge of sources 70 times, and where the information became available 13 times.

### *Discussion*

#### *Implications for Administrators*

Several issues having implications for school administrators appear to arise from these findings, even though they are based upon a small population of elementary teachers from one system. Teachers in this study viewed each other as productive sources of consultation and interacted regularly on professional matters which tended to focus on practical concerns and instructional procedures. With the exception of interpretation of curriculum guides these concerns arose from time to time throughout the school year. Teachers apparently are prepared to consult a number of their colleagues, presumably those whom they can conveniently contact and in whom they have some confidence. Apparently they regard specially designated teachers such as the principal, vice-principal and librarian in much the same light as they do other teachers. The level of satisfaction with consultative advice acquired in this way is generally high, although some dissatisfaction was reported. To what extent should this resource be further developed and formalized in order to take full advantage of it? Not all teachers consult in depth with their colleagues. Would more do so if more formal procedures were established and recognition given to this form of service? Or would an attempt to formalize a spontaneous consultative network lead to less use of peer consultation? Until more is known about teacher-teacher consultative patterns, each principal should develop a consultation structure appropriate to his school.

Consultative resources within school staffs probably will alter from year to year because of the ad hoc nature of the composition of staffs of most schools, annual staff changes, and in particular in-service programs they have experienced. Teachers' needs and potentialities within each school will also vary. The principal therefore continuously needs to assess his staff and each year, with their consent, develop an inventory of interests and talents which can be utilized. At the same time he can anticipate that certain regular needs will surface, such as curriculum advice, near the beginning of the school year, and advice on special pupil problems during the year. By taken action to meet these concerns through such an inventory, the most common obstacles for not obtaining help, namely lack of resources or lack of information on whom to consult, could be reduced.

A plausible explanation for the finding that the more experienced teachers were less satisfied with peer consultation than were their less



experienced colleagues is that the latter group lack the specific practical information which is needed to function in the classroom, but which is not provided for in preparation programs. Further, formal training raises the level of awareness of information needs and also sharpens the critical appraisal of any help which is forthcoming. A hypothetical relationship of this type merits further investigation perhaps leading to a theory of differential supervision in which, for instance, better-qualified experienced teachers are encouraged to act as resource teachers to less-qualified inexperienced teachers. At the same time the former group could be given easier access to specialized consultants at the system level, thus making more efficient use of human resources.

Other possible inferences are suggested. Those teachers with less education and more experience who are expressing less need for consultation, and who are therefore less likely to seek help, might profitably be given more opportunities to compare their teaching with that of others. However, strategies such as this based on restricted research data can be advocated only with caution.

Implications also arise for the system administrator. System consultants who have generalist training may feel most comfortable in general supervision, such as visits to new teachers. But, if teachers can get this kind of help directly from their in-school colleagues, then system-level consultation should probably emphasize more specialist services such as diagnosing learning difficulties. A move in this direction could have the double benefit of making school staffs more professionally responsible through increased collegiality and at the same time provide an incentive for further professional specialization of consultants.

### *Implications for Further Research*

Haughey's study was limited to three elementary schools in one school system which had very restricted central office consultative assistance available to teachers. Replication of this study in the following situations is highly recommended:

- (1) in secondary schools in systems which have few central office consultative staff;
- (2) in schools in systems which have substantial numbers of central office consultative staff;
- (3) in schools in sparsely populated systems which require considerable travel of central office staff; and
- (4) in schools in systems which utilize consultants based in "families of schools."

Data from such additional studies would again be largely descriptive and perceptual. Introduction of evaluative elements in researching this topic, although desirable, would be difficult. Participant-observer studies of the nature and effectiveness of different types of consultation in different settings (school level, number of consultants, location of consultants, etc.) may be possible. More attention needs to be given to statistical relationships among measures of expressed need and satisfaction and such variables as training and years of experience.

Now that many school districts across Canada are facing declining populations, and consequently have declining funds available for payment of personnel who do not instruct in a classroom, information about how teachers receive help with their difficulties would appear to be even more important than ever.

#### References

- Franc, L. H. Problems perceived by seventy-five beginning elementary school teachers. Unpublished doctoral dissertation, Columbia University, 1970.
- Fuller, F. F. Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 1969, 6, 207-226.
- Fuller, F. F., & Parsons, J. S. *Current research on the concerns of teachers*. Austin: Research and Development Centre for Teacher Education, Texas University, 1968.
- Haughey, M. L. Consultative practices in elementary schools. Unpublished doctoral dissertation, University of Alberta, 1976.
- Holdaway, E. A. Research into consultation. Presented at C.E.A. Short Course on Educational Leadership, Banff, Alberta, 1971.
- Kass, H. A concern-based developmental sequence of teacher professional growth. presented at the annual meeting of the Canadian Society for the Study of Education, Edmonton, Alberta, 1975.
- Lucio, W. H. The supervisory function: Overview, analysis, propositions. In W. H. Lucio (Ed.), *Supervision: Perspectives and propositions*. Washington, D.C.: Association for Supervision and Curriculum Development, 1967.
- MacKay, D. A. A systematic approach to consultation. Presented at the C.E.A. Short Course on Educational Leadership, Banff, Alberta, 1971.
- McGillivray, W. R. A survey of supervisory assistance as perceived by beginning teachers in selected urban high schools in Ontario. Unpublished master's thesis, University of Alberta, 1966.
- Miklos, E. Alternative approaches to supervision and consultation. Presented at the C.E.A. Short Course on Educational Leadership, Banff, Alberta, 1971.
- Milne, J. L. Problems of first year teachers in the urban school system. Unpublished master's thesis, University of Alberta, 1968.
- Moller, G. E. A comprehensive study of the problems of beginning teachers in selected large senior high schools. Unpublished doctoral dissertation, University of Nebraska, 1968.
- Parsons, G. L. Teacher perceptions of supervisory effectiveness. *The Canadian Administrator*, 1971, 11, 5-8.
- Plamondon, J. L. Consultative needs of teachers. Unpublished master's thesis, University of Alberta, 1973.
- Sergiovanni, T. J., and Starratt, R. J. *Emerging patterns of supervision: Human perspectives*. New York: McGraw-Hill, 1971.



THE JOURNAL OF  
**human resources**

EDUCATION, MANPOWER, AND WELFARE POLICIES

## Emphasizing Education—

costs of and returns from higher education and compensatory education programs;  
evaluations of vocational education, training, and vocational rehabilitation;  
discrimination in educational opportunities against blacks and other minority groups;  
production of educational outputs;  
supply of & demand for personnel in the education industry

*JHR* covers the broad spectrum of human resources including manpower policies, welfare policies, health care and education as they relate to the whole individual and his continuing role in society.

Regular issues of *The Journal of Human Resources* contain Articles, Review Articles, Communications, Books Received, and Book Reviews.

*in health care industry—*

*in manpower policies—*

*in education—*

*in welfare policies—*

human resources...

Special full-sized supplementary issues are published as time, material and funds become available. Previous areas covered include: "Vocational Education" and "Manpower Policies and Welfare Policies."

In 1974, The Journal of Human Resources provided the first published reports of the OEO Graduated Work Incentive Experiment, a 3-year field test of the effects on recipient families of 8 different "negative income tax" formulas.

## Subscribe to The Journal of Human Resources

*Founded in 1966, a quarterly publication*

*Editor: Stanley H. Masters*

*Subscription Rates: Institutions, \$20.00/year, 4 numbers; Individuals, \$10.00/year, 4 numbers (individuals must pre-pay)*

*Foreign postage: 50¢ per year additional*  
*Air Mail postage: \$12.50 per year additional*

**Send your subscriptions to:**  
**Journals Department**  
**University of Wisconsin Press**  
**P. O. Box 1379**  
**Madison, Wisconsin 53701**















## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centred and in capital letters, e.g., TABLE 4. The heading of the table is to be centred below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. the position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last name at the end of the manuscript under the heading, *References*.
6. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
7. In matters of style, the APA Publication Manual is considered definitive.





# ajer

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXIII

NUMBER 2

JUNE 1977

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA • EDMONTON

## THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

.....

Published quarterly in March, June, September and December  
by the  
Faculty of Education  
of  
The University of Alberta

.....

### Faculty Publications Committee

*K. A. Neufeld*, Department of Elementary Education  
*J. Chambers*, Department of Educational Psychology  
*A. K. Clark*, Department of Industrial and Vocational Education  
*A. G. Konrad*, Department of Educational Administration  
*M. R. Lupul*, Department of Educational Foundations  
*R. G. Martin*, Department of Secondary Education  
*C. H. Preitz*, Department of Industrial and Vocational Education

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

Chairman: *K. A. Neufeld*

Editor: *Andrew K. Clark*

Secretary: *Mrs. A. Onishenko*

.....

AJER gratefully acknowledges support from Canada Council and AACES. AJER is indexed in the *Canadian Education Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, and *Sociology of Education Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available in bound and unbound volumes; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, The University of Alberta, Edmonton, Alberta, Canada T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



# The Alberta Journal of Educational Research

Volume XXIII, No. 2

June, 1977

## CONTENTS

Change in Anxiety Level as a Factor in Test Performance .....	97
<i>J. B. Bond, Jr.</i>	
Practice Teaching Performance and Resultant Achievement Motivation .....	104
<i>F. B. W. Harper</i>	
Graduates' Assessment of Industrial Arts Teacher Education Programs .....	109
<i>C. J. Ross and L. D. Stewart</i>	
Conditional Reasoning in Adolescents .....	118
<i>L. C. Jansson</i>	
Differential Outcome Effects in Teacher Training Research as a Result of Functional versus Structural Behavioral Recording .....	128
<i>J. Martin and S. Auerbach</i>	
College Trustees View Their Selection .....	138
<i>A. G. Konrad</i>	
Knowledge and Revolution: The Student Movement in Tsarist Society (1856-1881) .....	151
<i>J. McLeish</i>	
Prediction of Metropolitan Readiness Test Scores .....	164
<i>E. A. Blowers</i>	
BOOK REVIEW	
<i>The Forked Road</i> , D. B. Creighton .....	169
Reviewed by H. T. Coutts	

FACULTY OF EDUCATION  
*The University of Alberta*





J. B. BOND, Jr.

*University of Manitoba*

## Change in Anxiety Level as a Factor in Test Performance

*The academic test performance of 110 female undergraduates was analyzed by test day anxiety, non-test day anxiety, and change in anxiety from a non-test day. The Today Form of the Multiple Affect Adjective Check List was administered on a non-test day and immediately prior to a test. Neither test day anxiety nor non-test day anxiety showed statistically reliable main effects in the analysis of test performance, but the interaction of the two was significant. Analysis of change in anxiety scores revealed that students who exhibited a high increase in anxiety on the test day performed more poorly on the test than students who had either a low increase or a high decrease in anxiety. The data support the view that an increase in anxiety exerts an interfering effect on test performance, and suggest that attempts be made to minimize increases and to decrease anxiety levels of students on test days. (Dr. Bond is Assistant Professor of Human Development in the Faculty of Home Economics, The University of Manitoba.)*

High anxiety states have been found to be related to poor test performance on a variety of tasks. High test-anxious persons take longer in the solution of anagrams (Russell & Sarason, 1965), have difficulty with probability learning tasks (Dusek & Hill, 1970), and receive lower grades on academic tests (Borr, 1972; Wine, 1971; Zuckerman, 1960). The performance of entire classes, as well as of individual students, has been found to be negatively related to anxiety levels (Osterhouse, 1975).

The effects of anxiety reduction approaches have not produced consistent effects on test performance (Cornish & Dilley, 1973; Mitchell & Ng, 1972; Smith, Ascough, Ettinger & Nelson, 1971). It is not surprising that counselling approaches have reported differing results, for the theoretical relationship between test anxiety and test performance is yet to be well established. There is some evidence that "anxiety" is not a unitary concept, affecting all persons similarly. High anxious females score more poorly than high anxious males (Russell & Sarason, 1965). In addition, the type of anxiety has been shown to be an important variable (Alpert & Haber, 1960).

Borr (1972) reported a correlation between debilitating anxiety and exam score but no relationship between facilitative anxiety and test score.

Finally, there have been differing reports concerning the relationship between level of anxiety and test performance. Some investigators have reported an inverted U-shaped relationship (Fein, 1963; Grooms & Endler, 1960; Sweeney, Smouse, Rupiper, & Munz, 1970). This position states that both low and high anxiety levels are related to poor test performance; optimal test performance is associated with moderate levels of test anxiety. An alternative position is that there is an inverse relationship between anxiety and test performance (Liebert & Morris, 1967; Osterhouse, 1975; Sarason, 1972; Wine, 1971).

For counsellors and teachers to control test anxiety, the relationship between test performance and anxiety needs to be more clearly delineated. One purpose of this study was to assess the relationship between test performance and the level of anxiety on a test day. If anxiety interferes with test performance, students with high levels of anxiety on a test day should perform less well than students with low levels of anxiety on the test day.

This study also compared test performance of students exhibiting various changes in anxiety levels between a non-test day and a test day. While anxiety levels generally increase on test days, some students may decrease in anxiety. It was hypothesized that students with high increases in anxiety would perform less well than students having a low increase, low decrease, or high decrease in anxiety, due to the interfering effects of anxiety (Borr, 1972; Osterhouse, 1975; Zuckerman & Lubin, 1965).

### *Method*

All students enrolled in a course in Family Development, 110 female Home Economics undergraduates, served as subjects. The Today Form of the Multiple Affect Adjective Check List (MAACL) was administered on two occasions, once on a non-test day, one month prior to a test, and once immediately prior to a test. The MAACL has been found to be an effective measure of anxiety (Zuckerman & Lubin, 1965). The anxiety score on the MAACL was determined for each student, and resulted in a test day anxiety score, a non-test day anxiety score, and a change in anxiety score (determined by subtracting test day anxiety from non-test day anxiety).

Following the collection of the data, median splits were used to form groups on the basis of both test day anxiety scores and non-test day anxiety scores. Anxiety measured on the test day produced a group of students ( $N = 59$ ) exhibiting high test day anxiety, with anxiety scores greater than nine, and another group of students ( $N = 51$ ), indicating low test day anxiety, with scores from zero to nine. Anxiety measured on the non-test day yielded a group of high non-test anxious students ( $N = 52$ ), with scores greater than eight, and a group of low non-test anxious students ( $N = 58$ ), whose scores ranged from zero to eight.

A second grouping procedure was used, by which the effects of change in anxiety from a non-test to a test day could be assessed. Two categories were first created on the basis of either an increase or a decrease in anxiety on the day of the test. A median split was then used for each category to establish four groups as indicated in Table 1. Eight students did not change in



TABLE 1  
GROUPS BASED ON CHANGE IN ANXIETY SCORES

Group	N	Range of Change in Anxiety Scores
High Increase	32	+6 to +18
Low Increase	33	+1 to + 5
Low Decrease	20	-1 to - 5
High Decrease	17	-6 to -17

anxiety level, and their test data were eliminated from further analysis.

The dependent variable in this study was performance on a multiple choice, true-false academic test of material presented during the previous month both through lecture and independent reading.

Results

Table 2 reports the mean test performance of groups formed on the basis of both test day and non-test day anxiety scores. Analysis of variance of this data (Table 3) revealed no statistically reliable main effects of either Test Day Anxiety or Non-test Day Anxiety. The interaction of Test Day Anxiety X Non-test Day Anxiety, however, was significant  $F(1,106)=3.655, p=.056$ .

TABLE 2  
MEAN TEST PERFORMANCE BY TEST AND NON-TEST DAY ANXIETY

		Test Day Anxiety	
		High	Low
Non-test Day Anxiety	High	61.79 (N=29)	60.52 (N=23)
	Low	58.53 (N=30)	61.10 (N=28)

TABLE 3  
ANOVA OF TEST PERFORMANCE BY TEST AND NON-TEST ANXIETY

Source	df	MS	F	Probability
Test Day Anxiety	1	16.099	.585	.999
Non-test Day Anxiety	1	60.443	2.197	.137
Test Day Anxiety × Non-test Day Anxiety	1	100.575	3.655	.056
Error	106	27.515		
Total	109	28.361		

The mean test performance of groups created on the basis of change in anxiety scores is shown in Table 4. A one-way analysis of variance (Table 5) indicated that change in anxiety was a significant element in test performance,  $F(3,98) = 3.625$ ,  $p = .024$ . To determine where the differences existed, a Least Significance Difference test was conducted (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). This test is a form of the Student's  $t$ , appropriate when an analysis of variance is significant, and a few groups are to be compared. A difference in mean test performance scores of 2.77 or greater was required for a significant difference at the  $p = .05$  level. The significant differences in test performance, therefore, were between High Increase and High Decrease, and High Increase and Low Increase. In both instances, students exhibiting a high increase performed less well on the test.

TABLE 4  
MEAN TEST PERFORMANCE BY CHANGE IN ANXIETY

Change in Anxiety	Test Performance
High Increase	58.125
Low Increase	61.939
Low Decrease	60.750
High Decrease	61.529

TABLE 5  
ANOVA OF TEST PERFORMANCE BY CHANGE IN ANXIETY

Source	$df$	MS	F	Probability
Between Groups	3	89.271	3.265	.024
Within Groups	98	27.341		
Total	101			

*Discussion*

As mentioned in the introduction, a conflict exists concerning how anxiety affects test performance. In this research, neither non-test day nor test day anxiety alone were found to affect test performance. The interaction of test day and non-test day anxiety, however, was significant. Perhaps the confusion resulting from previous research is due to a lack of measurement of baseline levels of anxiety on non-test days. For students who normally are anxious, high test day anxiety might only be reflective of a normal level of anxiety; however, for students who are generally low in anxiety level, high test day anxiety would indicate an abnormal level of anxiety. Similarly, students low in anxiety on both test and non-test days would be exhibiting a normal level of anxiety on the test day; students generally high in anxiety, but manifesting low levels of anxiety on test days, would be functioning at an abnormal anxiety level on the test day.

This interpretation is strengthened by the analysis of test performance by change in anxiety levels. When compared to students indicating a high increase in anxiety, those who had either a low increase or a high decrease in anxiety performed significantly better on the test. A high increase in anxiety, when compared to the normal anxiety state, seems to produce a lowered test performance.

This study has implications for both teachers and counsellors. The nonsignificant main effects of non-test and test day anxiety, and the importance of change in anxiety scores, emphasize the necessity of knowing and dealing with students as individuals, as well as members of a class. One cannot assume that poor academic test performance is due to a general effect of being in a high state of anxiety on the day of the test; the anxiety level of the individual must be compared to the level of anxiety on a non-test day. It appears to be appropriate to minimize anxiety increase on test days, for a high increase seems to result in lower performance, and high decreases in anxiety do not produce a parallel detrimental decrease in performance.

As a final note, this research contradicts, to some extent, the findings of Hountras, Grady and Vraa (1970), which reported differences between Americans and Canadians in the relationship between anxiety and academic achievement. Their study found that for Americans, a negative relationship existed between anxiety and performance on the American College Test, while for Canadians, a positive relationship existed. In addition, high anxiety Americans had the highest grade point average, while low anxiety Canadians had the lowest grade point average. While the authors did report that anxiety should not be taken as the cause of the differences in performance between these two groups, the opposite effects of anxiety in the two groups was, and still is, in need of further exploration. Certainly studies such as that of Hountras et al. (1970) indicate the need for further research on Canadian samples.

#### References

- Alpert, R., & Haber, R. N. Anxiety in academic achievement situations. *Journal of Abnormal and Social Psychology*, 1960, 61, 207-215.
- Borr, M. Relationship of test anxiety and academic performance when controlled for intelligence. *Journal of Clinical Psychology*, 1972, 28, 171-172.
- Cornish, R., & Dilley, J. Comparison of three methods of reducing test anxiety: Systematic desensitization, implosive therapy, and study counseling. *Journal of Counseling Psychology*, 1973, 20, 499-503.
- Dusek, J. B., & Hill, K. T. Probability learning as a function of sex of the subject, test anxiety, and percentage of reinforcement. *Developmental Psychology*, 1970, 3, 195-207.
- Fein, L. G. Evidence of a curvilinear relationship between IPAT anxiety and achievement at nursing school. *Journal of Clinical Psychology*, 1963, 10, 374-376.
- Grooms, R. R., & Endler, N. S. The effect of anxiety on academic achievement. *Journal of Educational Psychology*, 1960, 51, 299-304.
- Hountras, P. T., Grady, W. E., & Vraa, C. W. Manifest anxiety and academic achievement of American and Canadian college freshmen. *Journal of Psychology*, 1970, 76, 3-8.
- Liebert, R. M., & Morris, L. W. Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 1967, 20, 975-978.



- Mitchell, K. R., & Ng, K. T. Effects of group counseling and behavior therapy on academic achievement of test anxious students. *Journal of Counseling Psychology*, 1972, 19, 491-497.
- Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. *Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill, 1975.
- Osterhouse, R. A. Classroom anxiety and the examination performance of test-anxious students. *Journal of Educational Research*, 1975, 68, 247-249.
- Russell, D. G., & Sarason, I. G. Test anxiety, sex and experimental conditioning in relation to anagram solution. *Journal of Personality and Social Psychology*, 1965, 1, 493-496.
- Sarason, I. G. Experimental approaches to test anxiety: Attention and the uses of information. In C. R. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (Vol. 2). New York: Academic Press, 1972.
- Smith, R. E., Ascough, J., Ettinger, R. F., & Nelson, D. A. Humor, anxiety and task performance. *Journal of Personality and Social Psychology*, 1971, 19, 243-246.
- Sweeney, C. J., Smouse, A. D., Rupiper, O., & Munz, D. C. A test of the inverted-U hypothesis relating achievement anxiety and academic test performance. *Journal of Psychology*, 1970, 74, 267-273.
- Walter, D., Dentler, L., & Sarason, I. Anxiety and the intellectual performance of high school students. *Child Development*, 1964, 35, 917-926.
- Wine, J. Test anxiety and direction of attention. *Psychological Bulletin*, 1971, 76, 92-104.
- Zuckerman, M. The development of an affect adjective check list for the measurement of anxiety. *Journal of Consulting Psychology*. 1960, 24, 457-462.
- Zuckerman, M., & Lubin, B. *Manual for the Multiple Affect Adjective Check List*. San Diego: Educational and Industrial Testing Service, 1965.

# New Zealand Journal of Educational Studies

---

Published twice-yearly by the New Zealand Council for Educational Research

The New Zealand Journal of Educational Studies provides an avenue of publication for essays, research findings and critical comments in any of the major fields of educational study in New Zealand and the Pacific area.

The current subscription rate is US\$8.00.

Business correspondence and subscriptions should be addressed to the Executive Officer, New Zealand Council for Educational Research, Box 3237, Wellington, New Zealand.

## New Zealand Journal of Educational Studies

Some recent articles:

---

A Behaviour Modification Approach to Rehabilitating  
Behaviourally Disordered Children in an Adjustment Class  
*Lyn Fry and John Thomas*

Violent and Dangerous Acts on New Zealand Television  
*Stephen Ginpil*

Educational Reconstruction in Post-War Fiji: A New Zealand  
Challenge to British Colonial Policy *Clive Whitehead*

Unpublished Degree and Diploma Studies in Education, Part III  
*K.A. Pickens*

Educational Attainments of New Zealand's Managerial Elites  
*David S. Groatorex and Graham Fogelberg*

F. B. W. HARPER

*The University of Western Ontario*

## Practice Teaching Performance and Resultant Achievement Motivation

*Practice teaching is assumed to generate simultaneously two affective reactions, one the motive to succeed (MS) and the other the motive to avoid failure (MAF). The relative strengths of these two motives are predicted to be significantly related to the student teacher's performance in practice teaching. When MS is greater than MAF, the student teacher responds more positively and does better in the field experience than when MS is less than MAF. Two samples of men and women student teachers were assigned to the appropriate cell of a four-fold distribution of scores on a test of MS and a test of MAF. An analysis of variance of practice teaching marks showed that in the men's sample, the results confirmed the hypothesis, while in the women's sample, the hypothesis was partially confirmed. (Dr. Harper is Professor and Coordinator of Secondary Education at The University of Western Ontario in London.)*

Practice teaching generates at least two important affective reactions in student-teachers. For many students, practice teaching is positively viewed as a highly relevant component of the teacher preparation program, the opportunity to get into the practical or applied aspects of teaching and demonstrate their competence at the classroom level. From a theoretical point of view, the motive which seems appropriate to this attitude of positive commitment to school practice is the motive for success (MS). In achievement motivation theory, MS is defined as a relatively stable or enduring disposition to strive for success (Weiner, 1972). The second emotional reaction which is generated by practice teaching stems from the possibility of being found wanting in the skills appropriate to competent classroom teaching performance. The student teacher will be required in his practicum setting to perform under the critical scrutiny of the supervising teacher, his college professor, and the pupils. The student teacher may be apprehensive about failing in this important endeavor, and the negative emotional reactions associated with this apprehension give rise to a tendency to withdraw from or avoid the practice teaching experience. This



potential avoidant response is defined theoretically as the motive to avoid failure (MAF) (Weiner, 1972).

Both MS and MAF may take different values along a theoretical continuum, so that it is possible for a student teacher to be high, medium, or low in either scale according to specified criteria for determining the strengths of the motives. Atkinson (1964) has postulated that both MS and MAF may co-exist in the same person, and that the individual's behavior with respect to a defined goal, will be resultant of the strengths of the two motives, one acting in a positive goal-oriented direction, and the other in a direction away from the goal, as in an approach-avoidance conflict. Weiner (1972) has suggested that when MS is stronger than MAF, an individual's resultant behavior is positive, that is, he moves towards the goal. However, when MAF is stronger than MS, the individual is placed in a conflict position. His move towards the goal will be less positive and may in extreme cases be halted altogether. So far as practice teaching is concerned, a student teacher may fall into any one of a number of possible categories of resultant motivation stemming from the interaction of these two motives. He may be high in MS and low in MAF, in which case he would be expected from a theoretical point of view to regard practice teaching in a very positive light and be eager to go out and demonstrate his competency. On the other hand, he might be high in MAF and low in MS, and look forward with considerable apprehension to his practice teaching experience, perhaps, if his anxiety is high enough, not even turning up at all at his assigned school. In between are a number of theoretical positions, including the possibility of being high in both motives at the same time, or low in both at the same time.

The question arises as to whether the actual behavior of student teachers in the practice teaching situation is predictable from this kind of theoretical formulation. That is, will practice teaching performance, as reflected for example in the grades assigned to the student-teacher, differ as a function of the interaction of these two motives. Grades in practice teaching are obviously multiply determined, and the proportion of the variance that may be attributable to the effects of these motivational factors may be small. Nevertheless if there are significant differences associated with differences in motivation, control of these variables might be important for selection, placement, or remediation purposes.

The following hypothesis was tested. Student teachers' performance in practice teaching will be a function of the relative strengths of the two motives, MS and MAF. More specifically, those students rated as high in MS and low in MAF will receive higher grades in practice teaching than student teachers rated as low in MS and high in MAF.

Operationally, the measurement of motives such as MS and MAF has been the subject of considerable controversy. Atkinson's original work used the Thematic Apperception Test (TAT) as the measure of MS, and the Test Anxiety Questionnaire as the measure of MAF. Harper (1971, 1974), Weiner (1972), and Clarke (1973) have commented on the problems associated with these two operational measures. Clarke advocated the use of the N ach scale of the Personality Research Form (Jackson, 1967) in place of the TAT. Harper (1975) found that Jackson's N ach scale and the Debilitating Anxiety Scale (Deb) of the Achievement Scale (Alpert & Haber 1960)

significantly differentiated academic achievement in college students, when used as operational measures of MS and MAF respectively. Both of these scales have high reliability and the validation data presented by Harper suggested that the two scales are worthy of further study in achievement motivation research.

Method

Subject and Procedure

Two samples of student teachers were tested, one male and the other female. The students were enrolled in a post-baccalaureate program of teacher preparation for one year. There were 292 males and 299 females. Each subject was given the Personality Research Form and the Achievement Anxiety Scale at the beginning of the professional year. At the end of the year, the practice teaching grades were calculated as an average of the various marks assigned over the periods of the practicum experience.

Scoring and Analysis

The subjects were assigned to one of four cells in a 2x2 ANOVA design, according to their scores on the two scales, MS and MAF. Subjects scoring higher than the median on both scales were assigned to the High-High cell. Subjects scoring lower than the median in both scales were assigned to the Low-Low cell. Subjects scoring above the median on MS and below the median on MAF were assigned to the High-Low cell, and finally, subjects scoring low in the MS scale and high in the MAF scale were assigned to the Low-High cell. Subjects scoring at the median level on either scale were eliminated from the statistical analysis. The practice teaching marks were then entered into the ANOVA as the dependent variable, and the corresponding *F* values calculated as in a conventional 2x2 design.

TABLE 1  
PRACTICE TEACHING MARKS CLASSIFIED BY RESULTANT  
ACHIEVEMENT MOTIVATION

MEN		<u>Motive for Success</u>	
		High	Low
<u>Motive to Avoid Failure</u>	High	72.88	70.82
	Low	74.86	73.15
WOMEN		<u>Motive for Success</u>	
		High	Low
<u>Motive to Avoid Failure</u>	High	74.26	72.56
	Low	74.86	73.35

Results

The practice teaching mark averages for each cell of the 2x2 distribution are shown in Table 1 for each sex separately.

The high MS-low MAF cell had the highest average practice teaching mark, while the low MS-high MAF cell had the lowest average practice teaching mark in both sexes. This finding is consistent with the hypothesis.

A 2x2 ANOVA was performed on the data. The values for *F* are shown in Table 2. In the male sample, the values of *F* are significant for both main effects. In the female sample, the value of *F* for the effect due to differences in MS is significant, but the value of *F* for differences in MAF is not significant. In neither sample is there an interaction effect.

TABLE 2  
ANALYSIS OF VARIANCE OF PRACTICE TEACHING MARKS

Source	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	
MEN					
Between MAF	339.81	1	338.81	11.76	<i>p</i> < .01
Between MS	258.99	1	258.99	8.96	<i>p</i> < .01
Interaction	2.14	1	2.14		ns
Within Groups	8320.55	288	28.89		
Total	8921.49	291			
WOMEN					
Between MAF	39.62	1	39.62	1.79	ns
Between MS	212.49	1	212.49	9.61	<i>p</i> < .01
Interaction	0.78	1	0.78		ns
Within Groups	7162.56	296	22.11		
Total	7415.45	299			

Discussion

Practice teaching is a complex entity, and it is to be expected that many variables of both a personal and situational nature would interact to produce the final performance effects. It appears, from the results obtained in this study, that two of the variables which have significance for understanding performance in practice teaching are the motive for success and the motive to avoid failure. The motive for success in men and women is positively associated with getting higher grades. This finding accords with the general theory of achievement motivation that those who identify themselves as aspiring to accomplish difficult tasks and who respond positively to challenge and put forth effort to attain excellence are those who do, in fact, perform at a more effective level. Conversely, in men, the motive to avoid failure is negatively associated with getting higher grades in practice teaching. Again, this finding conforms to the theory of avoidant behavior, that those who are made anxious by competition and become



apprehensive and disabled by evaluative situations perform less well in demanding situations. Empirical testing of sample populations verifies, as Atkinson (1964) has suggested, that the two motives are often present to varying degrees in the same person at the same time. In practice teaching, the effects of the two motives are seen as a combination of their relative strengths. Student teachers who are rated as high in MS and low in MAF achieve significantly higher marks than those rated as low in MS and high in MAF. Those rated as intermediate in the two variables have practice teaching scores at an intermediate level. The hypothesis is thus sustained, though in the women's sample, the effects of the motive to avoid failure are less directly related to performance in practice teaching than is true for men. Further study of these variables with other samples will help determine whether this difference is consistent.

There are at least three possible implications of these findings for institutions which prepare teachers. First, the selection of candidates for admission may be assisted by measures of these motivational variables, along with more conventional criteria of an academic nature. Secondly, student teachers already admitted to teacher programs may be better matched to the field placements to which they are assigned if reliable and valid information about the relative strengths of these motives are known. Thirdly, remedial or therapeutic intervention with student-teachers who have potentially disabling motivational problems may be possible.

#### References

- Alpert, R. A., & Haber, R. N. Anxiety in academic achievement situations. *Journal of Abnormal and Social Psychology*, 1960, 61, 207-215.
- Atkinson, J. W. *An introduction to motivation*. Princeton, N.J.: D. Van Norstrand, 1964.
- Clarke, D. E. Measures of achievement and affiliation motivation. *Review of Educational Research*, 1973, 43, 41-51.
- Harper, F. B. W. Specific anxiety theory and the Mandler-Sarason Test Anxiety Questionnaire. *Educational and Psychological Measurement*, 1971, 31, 1011-1014.
- Harper, F. B. W. The comparative validity of the Mandler-Sarason Test Anxiety Questionnaire and the Achievement Anxiety Test. *Educational and Psychological Measurement*, 1974, 34, 961-966.
- Harper, F. B. W. The validity of some alternative measures of achievement motivation. *Educational and Psychological Measurement*, 1975, 35, 905-909.
- Jackson, D. N. *Personality research form manual*. New York: Research Psychologists Press Inc., 1967.
- Weiner, B. W. *Theories of motivation*. Chicago: Markham Publishing Company, 1972.

C. J. ROSS

and

L. D. STEWART

*The University of Alberta*

## Graduates' Assessment of Industrial Arts Teacher Education Programs

*This study involved the gathering of opinions held by Alberta teachers, who reported their perceived value of teacher education programs that had prepared them to teach industrial arts in Alberta secondary schools. A list of competencies needed in order to teach the Alberta Industrial Arts Curriculum in schools was developed and incorporated into a questionnaire which was distributed to all industrial arts teachers in the province.*

*As a result, the composition of the teaching force was determined, on the basis of location of teacher education program. The study found that the University of Alberta program was broader in scope than any other program investigated. However, University of Alberta graduates perceived their preparation as being adequate only to teach at the junior high school level.*

*Although teachers educated at institutions other than the University of Alberta perceived their preparation as being of relatively higher value, the data did not support the suggestion that such programs were in fact, of greater value. (Dr. Ross is a Sessional Lecturer with the Department of Industrial and Vocational Education, and Dr. Stewart is a Professor with the Department of Secondary Education at The University of Alberta.)*

The University of Alberta at Edmonton has been responsible for the preparation of industrial arts teachers since 1963, when it developed a program of teacher education designed to replace the industrial arts program then offered by the University of Alberta at Calgary. The thrust of the new program was to prepare teachers capable of implementing a technology based industrial arts curriculum. That program has operated for over a decade with little major revision.

The major purpose of this study was to determine the extent to which each Alberta industrial arts teacher perceived the value of his particular program of teacher education in developing the competencies needed to teach industrial arts in Alberta secondary schools. Irrespective of the

location of his teacher education, each Alberta industrial arts teacher is expected to teach the content specified by the provincial curriculum. However, there is no legal requirement for the University of Alberta to ensure that their program graduates do possess the competencies needed to teach that curriculum. For the sake of uniformity it was decided to accept the Alberta Department of Education curriculum as the authority on industrial arts content.

### *Method*

The study included: (1) the development of an instrument, listing the competencies needed to teach industrial arts; (2) the identification of the industrial arts teaching force on the basis of preparation program; (3) and the gathering of opinions held by teachers towards the adequacy of such programs.

### *Instrument*

The survey instrument was a questionnaire which incorporated a 6-point Likert-type scale. In order to maximize instrument validity the content was selected from the Alberta Department of Education curriculum for industrial arts, and validated by a panel of expert judges. Instrument reliability was maximized by the use of a large sample, in this case, the entire target population.

Another factor to be decided was the format of the response scale. The literature suggests that little noticeable difference in results occurs when either a multi-step scale or a dichotomous scale is employed. Research by Peabody (1962) found that bipolar scaled items could justifiably be scored dichotomously, dependent on the direction of the response, while Helmstader (1964) stated that Likert type scales enable "subjects to place themselves on an attitude continuum for each statement" (p. 133). It was decided to use a six-point bipolar scale (no value, most inadequate, inadequate, adequate, most adequate, excellent) for each content question, in order to force responses which reflected the direction of inadequacy or adequacy. However, as the study was concerned with perceptions of adequacy, it was further decided to collapse the data received into a dichotomous "inadequate/adequate" scale, because as Matell and Jacoby (1971) found, the conversion of multi-stepped scales to dichotomous measures did not significantly reduce either reliability or validity.

### *Competencies*

The questionnaire was designed to elicit the specialized competencies needed to teach the Industrial Arts Curriculum in Alberta schools. As no list of industrial arts teacher competencies was available, a set was developed in the following manner.

Reference to the Alberta Department of Education Program of Studies for High School Industrial Arts (1972), the Curriculum Guide for Industrial Arts General 10-20-30 (1968), and the Junior High School Curriculum Guide for Industrial Arts (1968), led to the identification of sixteen industrial arts content areas and the specific abilities that a school student may reasonably be expected to acquire after completion of a course of study in each of the



TABLE 1  
TABLE OF CONTENT AREAS AND COMPETENCIES

<u>Materials Areas</u>	<u>Materials Competencies</u>
Woods	Identification
Metals	Testing
Plastics	Hand tool use
Earths/ceramics	Machine tool use
Art metal craft	Fabrication
Leather craft	Safety procedures
Lapidary craft	Tool use comparison
Hot Metals	Tool maintenance
	Fastening methods
	Finishing methods
	Career guidance
	Instructional organization
<u>Technology Areas</u>	<u>Technology Competencies</u>
Photography	Principles
Printing	Comparison
Electricity	Component identification
Electronics	Component usage
Power Sources	Systems identification
Mechanical Transmission	Systems usage
Fluid Transmission	System assembly
Drafting	System operation
	System analysis
	Safety procedures
	Career guidance
	Instructional organization
<u>General Teaching Competencies</u>	
Apply adolescent psychology theory	
Communicate with adolescents	
Motivate and/or discipline students	
Teach both individuals and small groups	
Remedy unsafe laboratory conditions	
Handle student injury situations	
Develop a laboratory safety program	
Develop daily instructional plans	
Develop yearly instructional plans	
Inventory supplies and equipment	
Order supplies and equipment	
Budget supplies and equipment	
Organize student personnel	
Construct evaluative tests	

sixteen industrial arts content areas. Eleven skills common to the materials areas and 11 skills common to the technologies areas were identified. As it seemed logical that a successful industrial arts teacher should possess (at least) the same skills, each was redefined in the format of an industrial arts teacher competency. A further common competency was developed which dealt with the ability to organize and plan to teach each of the sixteen industrial arts content areas. Therefore, 192 competencies were identified in this manner (sixteen areas by twelve competencies each).

Another set of fourteen competencies was developed, which related to general teaching skills. By these means a total of 206 competencies were listed, as in Table 1, (192 specialized plus 14 general), which were typical of

those required of an industrial arts teacher in Alberta. The 206 competencies were representative of the sixteen industrial arts areas plus general teaching abilities, i.e. a total of seventeen industrial arts content areas.

*Population and Sample*

The population and sample consisted of all 414 industrial arts teachers in Alberta secondary schools during the 1975/76 school year. Teachers were contacted by mail and responses received from 331 or 80% of the population.

*Results*

*Composition*

The composition of the industrial arts teaching force, on the basis of location of teacher preparation is given in Table 2. Most teachers (64%) were prepared in Alberta, nearly one half (46%) having attended the University of Alberta program, while 18% were graduates of the superseded University of Alberta at Calgary program.

TABLE 2  
QUESTIONNAIRE RETURN SUMMARY

Teacher Education Institution	<i>n</i>	% of Population	% of Returns
University of Alberta	150	36	46
University of Alberta at Calgary	60	15	18
Other Canadian Institutions	14	3	4
U.S.A. Institutions	37	9	11
Other Institutions	8	2	2
Vocational Programs	16	4	5
Non-Industrial Arts Programs	17	4	5
Invalid Returns	29	7	9
Questionnaires Returned	331	80%	
Questionnaires Distributed	414	100%	

Of the remainder, 11% were prepared by various institutions in the United States of America, while 6% had attended unspecified Canadian institutions. It was also found that 10% of the teachers had received no formal industrial arts preparation at all, although half of these teachers held vocational qualifications which permitted them to teach industrial arts in Alberta schools. Nine percent of the responses were invalid.

*Program Assessment*

Before completing the main portion of the questionnaire, which related to the individual industrial arts competencies, teachers assessed the adequacy of their program as general preparation to teach industrial arts to Alberta secondary school students.

TABLE 3  
PERCENTAGE OF ALBERTA INDUSTRIAL ARTS TEACHERS WHO  
PERCEIVE THEIR PREPARATION AS ADEQUATE

Preparation Institution	Junior High	Senior High	Total	(n)
University of Alberta	69	36	46	(150)
University of Alberta at Calgary	78	68	18	(60)
Other Canadian Institutions	64	50	4	(14)
U.S.A. Institutions	97	95	11	(37)
Other Institutions	100	86	2	( 8)
Discarded returns			19	(62)

Table 3 discloses that 78% of the 60 University of Alberta at Calgary prepared teachers rated as adequate their preparation to teach industrial arts to junior high school students, as compared to 69% of the 150 University of Alberta prepared teachers. Between 64% and 100% of the 59 non-Alberta prepared teachers perceived as adequate their preparation to teach industrial arts to junior high school students. However, it should be noted that these non-Alberta figures refer to small samples and a wide variety of unspecified teacher education programs.

With regard to senior high school industrial arts preparation, 68% of the University of Alberta at Calgary prepared teachers felt adequately prepared to teach industrial arts to senior high school students, as compared to 36% of the University of Alberta prepared teachers. Between 50% and 95% of non-Alberta prepared teachers felt adequately prepared to teach industrial arts to senior high school students.

Table 4 refers to all Alberta industrial arts teachers, on the basis of teacher education program, and comprises the means of their perception of adequacy of preparation in each of the seventeen industrial arts content areas.

The table provides an overview of the types of programs, based on the opinions of Alberta industrial arts teachers. An examination of column A (University of Alberta), indicates that the program's emphasis was perceived to be towards four of the materials content areas (Woods, Metals, Plastics, Earths/ceramics), and all eight technology areas (Photography, Printing, Electricity, Electronics, Power Sources, Mechanical Transmission, Fluid Transmission, Drafting). The major perceived strength of the program was in preparation for Photography (60%).

Reference to column B (University of Alberta at Calgary), discloses that the program's emphasis was perceived to have been towards four of the materials content areas (Woods, Metals, Art Metal, Hot Metals), and four of the technology areas (Electricity, Power Sources, Mechanical Transmission, Drafting). The major perceived strength of the program was in preparation for Woods (86%), Drafting (84%), Metals (72%), Electricity (72%), and Power Sources (49%).

It appears that the major difference between these two programs is the



TABLE 4  
MEANS OF CONTENT AREA COMPETENCIES BY PREPARATION PROGRAM

	(A)	(B)	(C)	(D)	(E)	Total
<u>General</u>						
Teaching	43	55	49	86	89	53
<u>Materials</u>						
Woods	48	86	46	91	86	62
Metals	39	72	72	82	100	56
Plastics	46	15	25	84	29	42
Earths/ceramics	44	20	0	17	67	32
Art Metal	8	34	9	58	100	23
Leather craft	3	17	13	77	0	19
Lapidary craft	2	9	14	46	0	9
Hot Metals	20	46	59	67	79	36
<u>Technologies</u>						
Photography	60	18	29	32	0	43
Printing	47	7	29	65	66	41
Electricity	47	72	61	73	77	57
Electronics	43	20	18	43	0	36
Power Sources	47	49	62	78	0	53
Mechanical Transmission	31	38	57	46	0	36
Fluid Transmission	48	12	55	18	23	36
Drafting	35	84	74	88	93	56
Means	36%	37%	37%	59%	50%	41%
N	150	60	14	37	8	269

Note. Letters in parentheses indicate preparation institution(s):

- |  |                           |
|--|---------------------------|
| (A) = University of Alberta            | (D) = U.S.A. Institutions |
| (B) = University of Alberta at Calgary | (E) = Other Institutions  |
| (c) = Other Canadian Institutions      |                           |

added emphasis placed by the University of Alberta on technology preparation, as evidenced by the inclusion of the content areas of Photography, Printing, Electronics and Fluid Transmission, none of which formed part of the superseded University of Alberta at Calgary program. The University of Alberta apparently attempted to develop a broadly based program which was perceived by its graduates as being generally inadequate as preparation in all but one of the industrial arts content areas. On the other hand, the University of Alberta at Calgary program which was less technology-oriented, was perceived by its graduates to provide adequate preparation for five of the industrial arts content areas.

When referring to the remainder of the table, it is noted that columns C, D, and E do not pertain to any identifiable program, but are based on geographical areas. Those 14 teachers prepared in provinces other than Alberta (column C) appear to have experienced programs with a composite emphasis towards four of the materials areas (Woods, Metals, Plastics, Hot Metals), and seven of the technology areas (Photography, Printing, Electricity, Power Sources, Mechanical Transmission, Fluid Transmission, Drafting). Similarly, teachers prepared by various programs in the United

States of America (column C) appear to have experienced programs with a composite emphasis on seven of the materials areas (Woods, Metals, Plastics, Art Metal, Leather craft, Lapidary craft, Hot Metals), and seven of the technology areas (Photography, Printing, Electricity, Electronics, Power Sources, Mechanical Transmission, Drafting).

Column C refers to eight teachers prepared outside Canada or the United States of America, and was not relevant to the study.

### *Discussion*

The study disclosed that the University of Alberta program was perceived as providing preparation for more of the industrial arts content areas than did other preparatory programs which were investigated. Data were obtained which can identify specific competencies for which preparation at the University of Alberta was perceived as being either adequate or inadequate. By matching the perceived weaknesses against the assumed program objectives, specific competencies which appear to have been covered in insufficient depth may be located. Such information could then be used to further develop the University of Alberta program.

The results of the study have also provided base line data establishing the perceived worth of the various programs which have prepared Alberta teachers. It is difficult to make judgments about the value of the University of Alberta program in relation to the other programs. Because the research originated in Edmonton, it is possible that the responses of the University of Alberta program graduates may have been unusually critical and candid, due to a desire to provide evaluative feedback which could be used to improve that program. Conversely, graduates of non-University of Alberta programs may have been less critical (or more protective) of their particular teacher education program. Therefore, the finding that the perceived adequacy mean of the University of Alberta program was 36%, as compared to 37% and 59% for the University of Alberta at Calgary and the United States of America programs respectively, must be treated with caution. That is, it does not necessarily follow that programs which prepared industrial arts teachers in the United States of America were superior to the University of Alberta program.

Although the same questionnaire was used to investigate the various preparatory programs, the possible effects of the following factors should be considered:

1. The University of Alberta program had been in operation for less than ten years, while the University of Alberta at Calgary program ceased ten years ago. Therefore, University of Alberta graduates had taught for fewer years than had University of Alberta at Calgary graduates. The relationship between experience and opinion of program worth was not examined, as a major part of the study was concerned with the University of Alberta program. However, it is recognized that the teaching experience of non-Alberta educated teachers may range over an extended time period, which may have affected their perceptions.

2. Because the study originated in Edmonton, it is possible that University of Alberta graduates may have been more candid in their opinions than the graduates of other programs.



3. With the exception of teachers from either the University of Alberta or the University of Alberta at Calgary, who represent 46% and 18% of the sample respectively, the actual number of teachers from other educational institutions was small, (i.e. Other Canadian Institutions 4%, U.S.A. Institutions 11%, Other Institutions 2%), thereby making comparisons unreliable.

In the case of the University of Alberta, it becomes evident that the program developed along a broad base, possibly at the cost of depth in preparation, for graduates viewed themselves as being inadequately prepared to teach industrial arts to senior high school students.

The University of Alberta at Calgary program which began to phase out in 1965, apparently provided in-depth preparation in the basic areas of Woods, Metals, Drafting and Electricity. Program graduates viewed themselves as being adequately prepared to teach industrial arts to both junior high and senior high school students. It may be noted that those teachers had taught for ten or more years, consequently their assessment of that program could have been affected by the expertise which they subsequently may have acquired through teaching, although there is no evidence to support this opinion.

The viewpoint of teachers educated in the United States of America is interesting. Those teachers experienced a variety of educational programs and generally each held a very positive perception of his program's value, both in technical skill development and in professional preparation. Table 4 disclosed that those teachers perceived their preparation in General Teaching competencies (i.e. the professional component) to be of considerably more value than did teachers educated in Canada. Two possible explanations are likely. Firstly, it may be that programs in the United States of America are in fact better able to prepare industrial arts teachers for Alberta schools than any Canadian program. Secondly, the effect of being prepared in a country other than Canada may be that teachers exert more than the usual amount of effort in developing an industrial arts program in an Alberta school. Subsequently, it may appear that teaching in Alberta was less difficult than had been anticipated. Therefore, the teacher could form the impression that his preparation was superior, and the major reason for his effective teaching.

These suppositions could also explain the generally high ratings given by non-Canadian educated teachers, in comparison to the lower ratings given by Canadian educated industrial arts teachers. However, no data has been obtained to support either viewpoint.

### *Conclusions*

Although the University of Alberta program appears to be broader in scope than other programs investigated, its graduates generally perceived themselves as being less adequately prepared to teach industrial arts in Alberta than did graduates of other programs. However, the preparation of industrial arts teachers in Alberta is an area which requires more research and development. This study examined only one aspect of that complex subject, namely, the extent to which industrial arts teachers in Alberta



perceived the value of their particular program of preparation in developing selected competencies.

Nevertheless, the results of the study have indicated some features of the University of Alberta program which require attention. Consequently, it is recommended that follow-up studies be conducted to monitor the perceived worth of that program.

This study developed and tested a method of curriculum assessment which may be a useful aid in evaluating an ongoing program of teacher education. The method yields formative data which may lead to subsequent revision and improvement of such a program.

In addition to providing the basis of the research instrument, the set of industrial arts competencies may be useful to both practising and potential industrial arts teachers as a check list, to identify deficiencies in their own preparation which may require upgrading.

#### *References*

- Cochran, L. H. *Innovative programs in industrial education*. Bloomington, Illinois: McKnight and McKnight Publishing Company, 1970.
- Curriculum Guide for Industrial Arts General 10-20-30. Department of Education, Edmonton, 1968.
- Helmstader, G. C. *Principles of psychological measurement*. New York: Appleton Century-Crofts, 1964.
- Junior High School Curriculum Guide for Industrial Arts*. Department of Education, Edmonton, 1964.
- Junior High School Curriculum Guide for Industrial Arts*. Department of Education, Edmonton, 1969.
- Matell, M. S., & Jacoby, J. Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and Psychological Measurement*, 1971, 31, 657-674.
- Peabody, D. Two components in bipolar scales: Direction and extremeness. *Psychological Review*, 1962, 69, 65-73.
- Provus, M. Evaluation of ongoing programs in the public school system. In R. Tyler (Ed.), *Educational evaluation: New roles, new means*. Chicago: University of Chicago Press, 1969.
- Program of Studies for Junior High Schools of Alberta*. Department of Education, Edmonton, 1963.
- Program of Studies for Senior High Schools of Alberta*. Department of Education, Edmonton, 1972.
- Ross, C. J. An assessment of the Alberta Industrial Arts Teacher Education Program. Unpublished doctoral dissertation, University of Alberta, Edmonton, 1976.
- Stake, R. E. The countenance of educational evaluation. *Teachers' College Record*, 1967, 68, 523-540.
- Stufflebeam, D. L. The use and abuse of evaluation in Title III. *Theory Into Practice*, 1967, 6 (3), June, 126-133.
- Taylor, P. A. and Maguire, T. O. A theoretical evaluation model. *The Manitoba Journal of Educational Research*, 1966, 1, 12-17.

L. C. JANSSON

*University of Manitoba*

## Conditional Reasoning in Adolescents

*The purpose of this study was to assess conditional reasoning ability in adolescents as effected by the presence of suggestive and concrete familiar content and negations in various parts of the arguments. Four basic principles of reasoning were examined: modus ponens, inversion, conversion, contraposition. A total of 464 subjects in grades 8, 10, and 12 were given tests requiring a judgment of validity of simple conditional arguments presented in verbal form. Significant differences in performance were found among grade levels, content types, principles of reasoning, and modes (positions of a negation in the argument). Significant interactions were found between principle of reasoning and grade level, principle of reasoning and content type, and principle of reasoning and mode. (Dr. Jansson is Associate Professor in the Faculty of Education, The University of Manitoba.)*

Deductive reasoning, and in particular, conditional reasoning has become the focus of a number of investigations in recent years. Arguments of four basic types have frequently been considered:

1. Modus ponens (MP)
2. Inverse (INV)
3. Converse (CNV)
4. Contrapositive (CP)

These principles of deductive reasoning are well known in both symbolic and verbal form. Each argument has two premises and a conclusion. One premise must be implication, i.e., a conditional proposition of the form *if a, then b*, where *a* and *b* are simple propositions or statements. The second premise is the affirmation or negation of *a* or *b*, giving rise to the four argument types shown in Table 1. Conclusions follow necessarily in two of these four types, namely modus ponens and the contrapositive. In the other types no conclusion follows necessarily and Table 1 shows the usual nonvalid response. Test items used in research have generally asked subjects to decide whether or not a given conclusion follows necessarily from stated premises. Such research (e.g., Ennis & Paulus, 1965; O'Brien, 1972; Jansson, 1974a, 1975) has suggested that principles 1 and 4 cause little difficulty even for many elementary school children, while items involving

TABLE 1  
BASIC PRINCIPLES OF CONDITIONAL REASONING

Principle	Validity	Symbolic Form	Concrete Familiar Form	Suggestive Form
Modus Ponens (MP)	Yes	if $p$ , then $q$ $\frac{p}{q}$	If the dog is brown, then his name is Rover. The dog is brown. The dog's name is Rover.	If cars can whistle, then horses eat cheese. Cars can whistle. Horses eat cheese.
Inverse (INV)	No	if $p$ , then $q$ $\frac{\text{not } p}{\text{not } q}$	If the car in the garage is Mr. Smith's, then it is red. The car in the garage is not Mr. Smith's. The car is not red.	If rats are red, then snakes can jump. Rats are not red. Snakes cannot jump.
Converse (CNV)	No	if $p$ , then $q$ $\frac{q}{p}$	If the painting belongs to Don, then it is a painting of a horse. This is a painting of a horse. This painting belongs to Don.	If babies can read, then air is not fresh. Air is not fresh. Babies can read.
Contrapositive (CP)	Yes	if $p$ , then $q$ $\frac{\text{not } q}{\text{not } p}$	If the cat's name is not Tabby, then she is not white. The cat is white. The cat's name is Tabby.	If chickens cannot gallop, then dogs can chirp. Dogs cannot chirp. Chickens can gallop.

principles 2 and 3 are answered incorrectly, as judged by standard propositional logic, even by high school and university students.

Following the notation used in both standard logic as well as in the work of Piaget, in this paper  $p$  and  $q$  will stand for propositions, “.” for *and*; “V” for *or*; “ $\rightarrow$ ” for *implication*; and the bar or “-” for *negation*. Hence an argument may be symbolized, for example modus ponens becomes:  $p \rightarrow q$ ,  $p/q$ .

O'Brien (1972; O'Brien, Shapiro & Reali, 1971) has suggested that, given  $p \rightarrow q$ , subjects frequently construct the biconditional  $[(p.q)V(\bar{p}.\bar{q})]$  rather than the correct  $[(p.q)V(\bar{p}.q)V(\bar{p}.\bar{q})]$ . This behavior, termed “Child’s Logic” by O'Brien, is used with some consistency by subjects at all ages. Other researchers (Wilkins, 1928; Ennis & Paulus, 1965; Roberge, 1970) have focused on the semantic content of the propositions as a factor in their understanding, while Roberge (1969) and Eisenberg and McGinty (1974) have examined the effect of negations in different parts of the argument.

The present study is an examination of adolescents’ abilities to handle conditional reasoning as evidenced by their performances on test items of concrete familiar and suggestive content types. The results reported here are part of a larger investigation into differing testing procedures for conditional reasoning assessment. In this phase of the investigation, answers were sought to the following questions:

- 1) Do adolescents’ abilities to judge conditional arguments vary with chronological age (grade level)?
- 2) Do adolescents’ abilities to judge conditional arguments vary with the principle of reasoning?
- 3) Do adolescents’ abilities to judge conditional arguments vary with the type of content in the argument?
- 4) Do adolescents’ abilities to judge conditional arguments vary with the mode of the argument?
- 5) Are there significant interactions among age, principle of reasoning, content type, and mode?



6) What proportion of adolescents have mastered each of the four principles of reasoning?

The above questions refer to four important notions in this research. "Principles of Reasoning" has been covered in the introductory remarks while the following require further explication: content, mode, mastery.

*Content.* Content can be varied in numerous ways to determine its effect upon reasoning ability. A basic notion modified and used by others goes back to the work of Wilkins (1928), and more recently to the researches of Ennis and Paulus (1965), Ennis, Finkelstein, Smith and Wilson (1969), and Roberge (1970). The latter studies have used three content types: Concrete Familiar, Symbolic, and Suggestive. The present investigation restricted itself to the first and last of these three. Concrete Familiar content implies that vocabulary familiar to the subjects was used, and further, that the truth value of the premises was empirically neutral. Suggestive items, on the other hand, while retaining familiar vocabulary, contain at least part of one premise which is contrary to observable fact.

*Mode.* The term "mode" has been used in this study to identify forms of the implication  $p \rightarrow q$  which contain negations in either antecedant ( $p$ ) or consequent ( $q$ ). Such implications are used as the first premise in a simple conditional argument. Thus there are four modes:

- 1.  $p \rightarrow q$
- 2.  $\bar{p} \rightarrow q$
- 3.  $p \rightarrow \bar{q}$
- 4.  $\bar{p} \rightarrow \bar{q}$

where both  $p$  and  $q$  are considered to be propositions containing no negations. Given one of these modes of implication, the choice of proposition and parity for the second premise (i.e.,  $p$ ,  $\bar{p}$ ,  $q$ , or  $\bar{q}$ ) determines the principle of reasoning applying to the two-premise argument. Table 2 indicates these sixteen combinations, where  $p_1$ ,  $i = 1, 2$  is the premise of the argument. A number of studies mentioned have investigated what is here given by row 1, while, with a few notable exceptions, little has been reported for rows 2, 3, and 4. A systematic look at both dimensions is carried out in this investigation.

*Mastery.* Going beyond the mere examination of means and variances, there is an interest in determining whether or not a principle of logic has

TABLE 2  
NEGATIONS IN FIRST AND SECOND PREMISES OF A SIMPLE  
ARGUMENT AS THEY DETERMINE REASONING PRINCIPLE

<div>P<sub>1</sub> \ P<sub>2</sub></div>				
	$p$	$\bar{p}$	$q$	$\bar{q}$
1) $p \rightarrow q$	MP	INV	CNV	CP
2) $\bar{p} \rightarrow q$	INV	MP	CNV	CP
3) $p \rightarrow \bar{q}$	MP	INV	CP	CNV
4) $\bar{p} \rightarrow \bar{q}$	INV	MP	CP	CNV

been mastered by a subject. The choice of mastery criteria is not an empirical matter, but one which requires serious consideration if the criteria are to be satisfactory. This issue is given explicit attention in the work of the Cornell Critical Thinking Project (Ennis & Paulus, 1965, 14-18) and in such articles as those by Ennis (1967) and Broudy (1961) so that a lengthy discussion will not be presented here. Suffice it to say that the result of these considerations is an operational definition which has been used in a number of studies, including this one. Studies which have a similar logical frame of reference and operational definition are thus somewhat comparable. Specifically, "mastery" is defined as a stated proportion of items correct, e.g., 6 out of 8 items correct, where a principle subtest consists of eight items. Using cumulative binominal probabilities, it can be determined that the probability of showing mastery of a principle by chance alone is approximately .02 for the format used. Such a criterion, while admittedly arbitrary, seems "reasonable."

### *Method*

#### *Tests*

Two pencil and paper tests were used in this investigation:

1. *The Jansson Conditional Reasoning Test*, Form CF
2. *The Jansson Conditional Reasoning Test*, Form SG

These two tests contained conditional reasoning items of the two content types. An example from test form CF follows:

Suppose you know that

If the horse is black, then it is not named Flicka.

The horse is not named Flicka.

Then would this be true?

The horse is not black.

Three responses were possible and were defined as follows for the examinees:

1. YES      It must be true.
2. NO      It can't be true.
3. MAYBE It may be true or it may not be true. You weren't told enough to be certain whether it is "YES" or "NO".

Each test contained 32 items, 8 for each of four principles of conditional reasoning, viz. modus ponens, inverse, converse, and contrapositive. A systematic variation of negations in the first premises of the arguments was employed to construct differences in mode. A further factor in the testing was the use of concrete familiar (CF) and suggestive (SG) content in the items. Each test contained items of one type only, randomly arranged by principle and mode (the same order on both tests).

#### *Procedure*

The sample was drawn from grade 8, 10, and 12 mathematics students in four Winnipeg area schools during the winter of 1975. Students were assigned randomly to one of the two test forms, giving the breakdown shown in Table 3. Because of this random assignment, there is no reason to

TABLE 3  
NUMBERS OF STUDENTS IN SAMPLE BY GRADE LEVEL AND TEST FORM

Grade	Test Form		
	CF	SG	Combined
8	96	89	185
10	63	77	140
12	68	71	139
All Grades	227	237	464

suspect that the two test groups (CF and SG) are inherently any different with respect to their logical reasoning abilities.

In order to answer the questions posed, a 3x2x4x4 analysis of variance with repeated measures on the last two factors was performed (Winer, 1962, 298-353). The first factor was that of grade level, i.e., grade 8, 10, or 12. The second factor was that of content type (or test form). The third and fourth factors were those of principle of reasoning and mode, respectively. While post hoc comparisons of means were not carried out statistically, graphing of data showed trends and interactions. Further analyses indicated the proportions mastering each principle.

Results

The dependent variable in the analysis of variance was subjects' scores on the conditional reasoning tests. These can be analyzed in terms of principle (P) or mode (M) subtests; each subtest contains eight items and hence possible scores range from 0 to 8. The analysis of variance results are given in Table 4 and a cross tabulation of means is presented in Table 5 for all four factors. The number of test items in each cell of the latter table is 2, and hence the maximum cell score is 2.00. Table 6 gives means for each principle by grade level.

The significant ( $p < .001$ ) *F*-ratio for the factor grade level (L) considered together with the means for each level indicate significant improvement with age as students move from grade 8 to grade 12. For the factor content the significant ( $p < .001$ ) *F*-ratio suggests that concrete familiar content is handled more easily than suggestive content when other factors are not considered. No significant ( $p > .05$ ) interaction was detected for the factors content and grade level.

The significant ( $p < .001$ ) *F*-ratio for the factor principles of reasoning indicates that abilities varied significantly among the four principles investigated. The means in Table 6 further suggest that the order of difficulty of the principles, greatest to least, for all grades combined, is: converse, inverse, contrapositive, and modus ponens. This confirms earlier findings (Jansson, 1975) for preservice elementary school teachers. While significant interactions of principle and content and principle and level were shown to exist in this sample, in all relevant combinations the inverse proved to be easier than the converse. That is, the inverse was easier in both contents and at all three levels. This result, while not consistent with findings in all previous studies, does seem to confirm a trend. The graph in



TABLE 4  
ANALYSIS OF VARIANCE FOR GRADE LEVEL, CONTENT,  
PRINCIPLE AND MODE

Source	<i>df</i>	<i>MS</i>	<i>F</i>
<u>Between Subjects</u>			
L (Grade Level)	2	27.99	45.77**
C (Content)	1	4.15	6.79**
L × C	2	.52	.85
Subjects within Groups	458	.61	
<u>Within Subjects</u>			
P (Principle)	3	713.70	787.17**
P × L	6	6.80	7.50**
P × C	3	3.65	4.02*
P × L × C	6	.83	.91
P × Subjects within Groups	1374	.91	
M (Mode)	3	18.55	69.40**
M × L	6	2.20	8.22**
M × C	3	9.49	35.51**
M × L × C	6	1.55	5.78**
M × Subjects within Groups	1374	.27	
P × M	9	4.76	18.38**
P × M × L	18	.73	2.83**
P × M × C	9	4.30	16.59**
P × M × L × C	18	.43	1.67
P × M × Subjects within Groups	4122	.26	

\*  $p < .01$   
\*\*  $p < .001$

Figure 1 illustrates the LxP interaction. The grade 12 performance on the contrapositive has so far defied explanation. One might wonder if the instructions in administering the test might have been faulty. These, however, were standardized at all grade levels and schools. The grade 12 students came from two different schools in different school divisions and the LxP graph for the SG looks just like the one given in Figure 1 for CF. Thus, with the exception of the contrapositive, this investigation's results support and confirm those of others with respect to an increasing ability with age to judge deductive arguments.

Given the implication  $p \rightarrow q$ , the mode is determined by the locations of negations within the propositions  $p$  and  $q$ . While the factor mode had a significant ( $p < .001$ )  $F$ -ratio, the significant interactions suggest that no generalization can be readily made about an order of difficulty of modes, although there is generally increasing ability over the grade 8 to 12 range in all modes (except when the CP applies). The graphs in Figure 2 show the interactions of mode with level and content (CxLxM). Mode 3 (negation in

TABLE 5  
CROSS TABULATION OF MEANS FOR ANALYSIS OF VARIANCE:  
GRADE LEVEL, CONTENT, PRINCIPLE, AND MODE

Principle	Mode	CF			SG		
		8	10	12	8	10	12
1	1	1.59	1.65	1.78	1.63	1.70	1.90
1	2	1.45	1.56	1.63	1.52	1.56	1.75
1	3	1.53	1.75	1.85	1.61	1.64	1.86
1	4	1.49	1.35	1.69	1.30	1.45	1.63
2	1	.43	.73	.96	.27	.44	.92
2	2	.24	.11	.26	.24	.40	.58
2	3	.43	.84	1.12	.10	.14	.37
2	4	.38	.56	.91	.28	.36	.70
3	1	.17	.13	.29	.17	.13	.52
3	2	.18	.14	.16	.24	.26	.42
3	3	.39	.65	.76	.12	.32	.34
3	4	.23	.27	.44	.25	.29	.38
4	1	1.65	1.59	1.37	1.46	1.45	1.25
4	2	1.07	1.11	1.01	1.12	1.03	.76
4	3	1.08	1.48	1.46	1.22	1.30	1.28
4	4	1.96	1.21	1.28	1.46	1.58	1.32

TABLE 6  
CROSS TABULATION OF MEANS: GRADE LEVEL,  
CONTENT, AND PRINCIPLE

		Principles				
		MP	INV	CNV	CP	ALL
CF	8	6.06	1.47	.96	4.76	12.25
	10	6.30	2.24	1.19	5.38	15.11
	12	6.96	3.25	1.66	5.12	16.99
3 Grades		6.39	2.22	1.23	5.04	14.88
SG	8	6.05	.90	.78	5.27	13.00
	10	6.35	1.35	1.00	5.36	14.06
	12	7.14	2.56	1.66	4.62	15.99
3 Grades		6.47	1.54	1.11	5.11	14.23

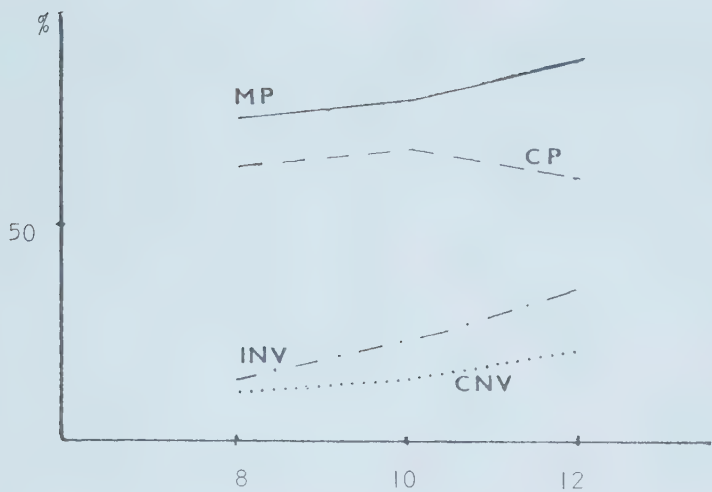
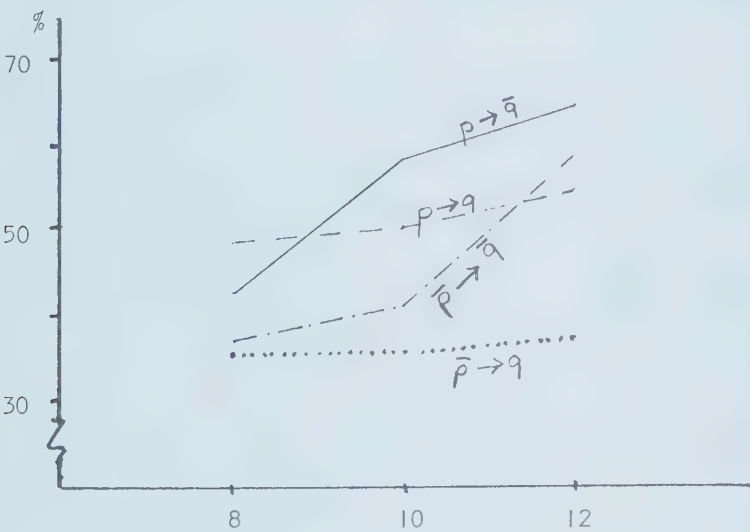
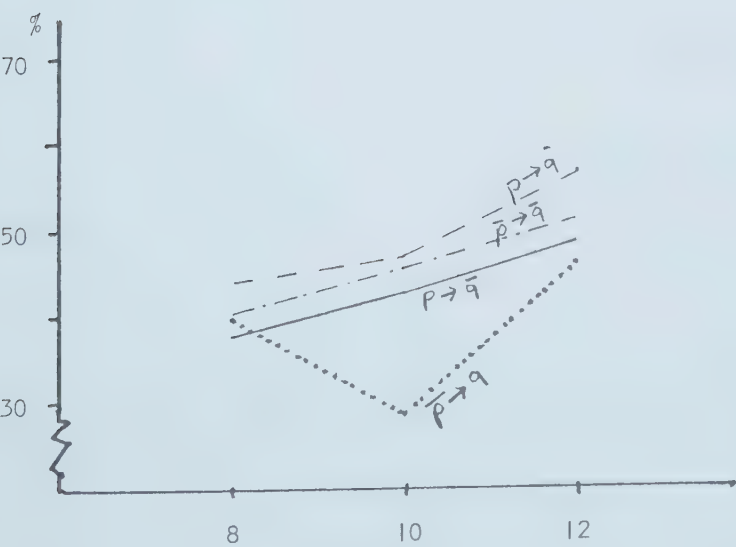


Figure 1. Grade level by principle of reasoning interaction (concrete familiar content only).



(a) Concrete Familiar



(b) Suggestive

Figure 2. Content by grade level by mode interactions.



the consequent of the first premise) was evidently more difficult in the suggestive form. Perhaps this was caused by subjects confusing it with the contrapositive form of argument. Of eight items of this type, two came from each of four principles and thus four items had as second premise some form of the consequent. Suggested by this analysis is a mode by principle interaction, which was in fact significant ( $p < .001$ ). Examination of the mode data reveals an extremely complex picture from which no easily describable pattern emerges. The breakdown into so many components, while perhaps providing an accurate description, may lead to such complexity as to be of little pragmatic value. A linear regression analysis reported earlier (Jansson, 1974b) suggested that, in fact, for this type of test the factor principle contributed significantly more to the variance in problem difficulty than all of the other variables combined. The results on modes confirm the findings of O'Brien (1972) not only in that  $\bar{p} \rightarrow q$  is the most difficult at all levels, but also that in grade 8 (9 in the O'Brien study)  $p \rightarrow q$  is easiest and by grade 12 (11)  $p \rightarrow \bar{q}$  was easiest, at least for concrete familiar items.

Proportions of subjects mastering each principle are shown in Table 7. While these proportions tend to run a little lower than those reported in some other studies (Ennis & Paulus, 1965; Roberge, 1970; Jansson, 1975), the general pattern of increasing mastery with age is again confirmed. The finding that the inverse is somewhat easier than the converse is repeated here also. Overall low scores are probably explainable in part by the relatively large number of negatives in the arguments of this test — fully three-fourths of the items contained a negation in the first premise.

TABLE 7  
PROPORTIONS OF SUBJECTS MASTERING EACH PRINCIPLE

Principle	Grade Level			
	8	10	12	Combined
MP	.71	.76	.93	.79
INV	.06	.05	.19	.10
CNV	.01	.03	.06	.03
CP	.43	.53	.44	.46

Discussion

In looking back at the questions to which this investigation sought answers, one finds that questions (1)-(5) may be answered in the affirmative, while data on (6) has been presented in Table 7. Not only the data in Table 7, but also that in Table 6 indicate a rather low level of performance on basic logical tasks. Data of this and similar sorts collected in a number of studies, many in recent years, make it abundantly clear that neither children nor adults think according to the canons of formal logic. To put it another way, formal logic does not serve as a very good model for human reasoning.

Piaget has attempted to use elements of logic to describe formal reasoning and its development (Inhelder & Piaget, 1958), and a number of interesting studies have attempted to confirm or refute Piaget's hypotheses. Between the time of inception of this study and its conclusion an in-depth conceptual critique of Piaget's logic was published by Ennis (1975). It suffices to say here that this latter study suggests that Piaget's logical model is not adequate with respect to either logic or its description of reasoning behavior. Hopefully studies of the sort reported here will enable researchers to document such behavior more precisely and to construct more adequate models of the development of logical reasoning.

Based in part on research carried out with a grant from the Research Grants Committee, Faculty of Graduate Studies, University of Manitoba.

### *References*

- Broudy, H. S. Mastery. In B. O. Smith & R. H. Ennis (Eds.), *Language and concepts in education*. Chicago: Rand McNally, 1961.
- Eisenberg, T. A., & McGinty, R. L. On comparing error patterns and the effect of maturation in a unit on sentential logic. *Journal for Research in Mathematics Education*, 1974, 5, 225-237.
- Ennis, R. H. Readiness to master a principle. In B. P. Komisar & C. J. B. Macmillan (Eds.), *Psychological concepts in education*. Chicago: Rand McNally, 1967.
- Ennis, R. H. Children's ability to handle Piaget's propositional logic: A conceptual critique. *Review of Educational Research*, 1975, 45, 1-41.
- Ennis, R. H., & Paulus, D. H. *Critical thinking readiness in grades 1-12. (Phase 1: Deductive reasoning in adolescence)*. Ithaca, N.Y.: Cornell Critical Thinking Project, 1965. (ERIC Document Reproduction Service No. ED 003 818)
- Ennis, R. H., Finkelstein, M. R., Smith, E. L., Wilson, N. H. *Conditional logic and children (Phase II C)*. Ithaca, N.Y.: Cornell Critical Thinking Project, 1969. (ERIC Document Reproduction Service No. ED 040 437)
- Inhelder, B., & Piaget, J. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books, 1958.
- Jansson, L. C. *The development of deductive reasoning: A review of the literature (Preliminary Version)*. Winnipeg, Canada: University of Manitoba, 1974. (ERIC Document Reproduction Service No. ED 090 034) (a)
- Jansson, L. C. Structural and linguistic variables that contribute to difficulty in the judgment of simple verbal deductive arguments. *Educational Studies in Mathematics*, 1974, 5, 493-505. (b)
- Jansson, L. C. The judgment of simple deductive arguments by pre-service elementary school teachers. *Alberta Journal of Educational Research*, 1975, 21, (1), 1-10.
- O'Brien, T. C. Logical thinking in adolescents. *Educational Studies in Mathematics*, 1974, 4, 401-428.
- O'Brien, T. C., Shapiro, B. J., & Reali, N. C. Logical thinking—Language and context. *Educational Studies in Mathematics*, 1971, 4, 201-219.
- Roberge, J. J. Negation in the major premise as a factor in children's deductive reasoning. *School Science and Mathematics*, 1969, 69, 715-722.
- Roberge, J. J. A study of children's abilities to reason with basic principles of deductive reasoning. *American Educational Research Journal*, 1970, 7, 583-596.
- Wilkins, M. C. The effect of changed material on ability to do formal syllogistic reasoning. *Archives of Psychology*, 1928, No. 102.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

J. MARTIN

and

S. AUERBACH

*Simon Fraser University*

## Differential Outcome Effects in Teacher Training Research as a Result of Functional versus Structural Behavioral Recording

*Teacher behavior variables in an experimental teacher training study were coded in two ways; one method of coding was based solely upon the formal structural properties of discrete teaching acts, the other method of coding was based upon the functional properties (consequential student behaviors) of discrete teaching acts. Differential outcomes on such variables as teacher questioning and reacting were associated with the different coding methods and units of analysis employed. Implications of the findings for research on teaching and teacher training programs were presented. (Dr. Martin is Assistant Professor in the Faculty of Education and Mr. Auerbach is a graduate student in the same faculty at Simon Fraser University.)*

Research on teacher training is generally taken to mean research in which teachers' behaviors and characteristics serve as dependent variables, and teacher education programs and procedures serve as independent variables (Gage, 1971). Teacher effectiveness research, on the other hand, takes teachers' behaviors and characteristics as independent variables and examines their effects upon the dependent variables of pupils' learning and behaviors (Gage, 1971). Together, research on teacher education and research on teacher effects can be considered to constitute research on teaching as it has been practised over the past two decades.

Research on teacher training has been quite consistent and clear in its findings. Three main classes of independent variables have received particular attention, and their effects in changing teacher behaviors are reasonably well documented. MacDonald (1969) refers to these variables as requisites for the acquisition of teaching skills, and describes them as: (1) specifying the skill precisely; (2) providing practice under carefully structured conditions; and (3) providing prompt and specific feedback. The



same general variables are referenced by Gage and Berliner (1975) as pre-performance factors, during performance factors, and post-performance factors.

At present, a wide variety of teacher-training products (e.g., Borg, Kelley, Langer & Gall, 1970; Houston, et al., 1973; Program on Teaching Effectiveness, 1974) are available which incorporate numerous procedures aimed at increasing the impact of the independent variables of specificity, practice, and feedback. These products attempt to focus the user's (teacher's) attention upon precisely defined teacher skills via observational learning or modelling (Perry, Leventhal, & Breen, 1975), provide focussed but non-threatening theatres for rehearsal (microteaching, minicourses, etc.), and provide immediate and exact feedback with the assistance of structured observation systems (Flanders, 1970; Simon & Boyer, 1970) and/or videotape recordings (Acheson, 1964). Combined feedback from structured observation instruments and VTR equipment has now become an integral part of those teacher training procedures which have successfully changed teacher behaviors in specified and "theoretically desirable" directions (cf. Berliner, 1969; Gage & Berliner, 1975).

While research on teacher training has produced a consistent body of information, teacher effectiveness research has been largely inconclusive and unproductive. Recent reviews on teacher effectiveness (Rosenshine, 1971; Dunkin & Biddle, 1974) have attempted to identify positive relationships between teacher variables and student learning. To date, research on teacher effectiveness has identified only a few broadly defined or high inference teacher behaviors that have consistently low to moderate relationships with student achievement. A pessimistic summary of what has been learned from this research is that almost anyone who is polite to students and who presents material in a common sense way would be approximately as effective as the most highly trained "superteacher" (Winne, 1976).

Thus, the overwhelming problem posed by the most recent information from research on teaching is that it is possible to train teachers to behave in specified ways, but it seems impossible to determine which of these ways contributes to the academic achievement and social learning of children in classrooms. This is indeed a rather sobering, if not totally embarrassing, paradox for the educational researcher.

While there may be many reasons contributing to this less than desirable state of affairs, one likely element is that definitions of teaching variables have not been carefully examined for the possibility that they have varying degrees of influence under different conditions (Winne, 1976). Training techniques based on the findings of teacher training research have tended to facilitate the acquisition of discrete structurally-defined teaching skills. They have not paid particular attention to the various contexts in which such skills are employed. For example, teachers might learn to ask a greater number of higher-order questions at the end of a teacher training program, but the "appropriateness" with which they employ such specific skills might not change. That a teacher asks more conceptual or theoretical questions does not necessarily mean that he (or she) elicits more conceptual or theoretical student responses. If he asks such questions when students have

not been prepared to compare, contrast and examine relevant information, or when they have not been attending to the prerequisite discussion, such student responses are unlikely, if not impossible.

Increases in the rates with which teachers emit discrete teaching acts which purport (on the basis of one theory or another) to facilitate student learning, may well be meaningless if they are not functionally tied to preceding and consequent instructional events. The use of structural skills without the ability to discriminate appropriate stimulus contexts in which such skills might be functionally embedded, and without the ability to examine whether the emission of such skills leads to intended student responses can be rather vacuous. Conceptual questions which result in factual student responses may be structurally conceptual, but are still factual in their function. The interactive changes in student responses during actual instruction can be viewed as equivalent to short-term student learning (Becker, Englemann, & Thomas, 1971). If the application of discrete teaching behaviors does not produce such short-term learning, how can it be expected to result in significant longer-term student learning outcomes? If it is true that teacher training techniques often result in structural rather than functional changes in teaching behavior, the pessimistic findings of teacher effectiveness research are quite understandable.

The present study considered both structural and functional changes in two crucial aspects of teacher performance (teacher soliciting and teacher reacting) as a consequence of four brief teacher training treatments. A structural teaching act was recorded on the basis of formal teacher behavior independent of consequent student responses. A functional teaching act was recorded on the basis of consequent student responses. While the experiment was formally a teacher training study, its consideration of short-term interactive changes in student performance provides interesting links to the more traditional teacher effectiveness experiments. The results of the teacher training study, while interesting, must be interpreted cautiously since the small sample size, and resulting low levels of power in statistical tests, makes generalization extremely tenuous. Differential outcomes as a consequence of structural versus functional recording of teacher behaviors form the focus of the report. The experiment was thus predominately methodological in its orientation.

### *Method*

#### *Sample and Setting*

Participants in the experiment were twelve students enrolled in the Simon Fraser University Professional Development Program. This program covers twelve months and is divided into three four-month blocks. During the five weeks of the study, the twelve students were in the second of these trimesters — a four-month in-school practicum. All subjects had been placed in elementary school classrooms (grades 1-6) in an urban centre. Their practicum work was supervised by a single faculty associate — an experienced teacher who, along with many others, had been hired by the university to provide this service. The twelve classrooms in which subjects had been placed were randomly assigned to one of four experimental conditions.



## *Differential Outcome Effects in Teacher Training*

Subjects ranged in age from 19 to 26 years. Only one of the twelve had previous classroom teaching experience. The twelve classrooms were located in schools which were part of the public school system in the province of British Columbia. All offered traditional department of education curricula and presented few unique features in the way of open-areas, innovative teaching methods, special student populations, etc.

### *Dependent Variables and Measures*

The eight dependent variables considered are defined in Table 1. The recording of all dependent variables was done from videotapes of the actual teaching sequences. Trained observers recorded numerical values (in the case of the talk variables) or frequencies (in the case of the questioning and reacting variables) on standard scoring grids consisting of matrices of interlocking squares in which rows represented different variables, and movement across columns yielded a pattern of behaviors over time. Three separate grids were employed for each of the variable classes — talk, questioning, and reacting. At the end of the recording periods, the total frequency (or numerical total) of each dependent variable was compared with the total frequency (or numerical total) for its generic class, and percentages were determined.

TABLE 1  
DEFINITIONS OF DEPENDENT VARIABLES

Variable	Operational Definition
Quantity of Teacher Talk	One word remarks = 1 False starts (incomplete sentences) = 3 Complete sentences = 5
Quantity of Student Talk	One word remarks = 1 False starts (incomplete sentences) = 3 Complete sentences = 5
Factual Questions	Teacher questions which require pupils to recall isolated facts
Conceptual Questions	Teacher questions which require pupils to put facts together in explaining, making comparisons and contrasts, giving relationships, etc.
Theoretical Questions	Teacher questions which require pupils to speculate and infer from facts they possess
Closing Reactions	Teacher reactions which require no further comment from the pupils to whom they are directed
Sustaining Reactions	Teacher reactions which invite pupils to continue speaking, but do nothing to enrich or elaborate upon what they are saying
Extending Reactions	Teacher reactions which invite pupils to continue speaking, but require them to engage in more complex thinking.



### *Experimental Treatments and Procedures*

After a fifteen-minute pretreatment recitation lesson by each of the twelve interns had been videotaped, the interns and their respective practicum classrooms were randomly assigned to one of four treatment conditions in a 2x2 factorial design. These conditions differed in terms of the feedback media employed to focus the interns' attention on their teaching performance in the areas of classroom talk, teacher soliciting, and teacher reacting.

The eight dependent variables described in the preceding section were precisely defined and modelled to all the interns in a sequential fashion over a four-week experimental period. Teacher-pupil talk was highlighted during the first week, teacher soliciting during the second, and teacher reacting during the third. The final week of the experiment was concerned with all three skill areas. The faculty associate met with each student once a week and defined and modelled the designated teacher skills during a half-hour preinstructional conference. This session was followed by a thirty-minute instructional period during which the interns practised the skills in a regular classroom setting while the faculty associate observed. All interns were told that the objectives of the four-week period were to increase the amount of pupil talk while decreasing their own teacher talk, to increase their output of conceptual and theoretical questions while decreasing their emission of factual questions, and to increase the number of sustaining and extending teacher reactions while reducing the frequency of closing reactions. Interns in all treatment conditions received an equivalent amount of skill specification and practice time.

The thirty-minute instructional periods during which the interns practised the teaching skills and attempted to change their relative emission frequencies were videotaped in two of the four treatment conditions. In two of the four treatment conditions, one which employed videotape and one which did not, the faculty associate recorded classroom talk and teacher questions and reactions on the standard scoring grids described in the preceding section. In this manner, one treatment group used neither videotape nor structured observation (C), a second used videotape alone (V), a third used structured observation alone (O), and a fourth group employed both videotape and structured observation (VO).

After each thirty-minute instructional period, the faculty associate spent some time analyzing the videotapes and/or structured observation schedules, or reflecting upon his general observations. In all conditions, this phase was followed by a forty-five-minute postinstructional feedback session between intern and faculty associate. These feedback sessions made use of videotapes and structured observation schedules where these were available. Feedback in group C was dependent solely upon recollection and reconstruction. The faculty associate attempted to employ the same level of precision in his expressive communications across all feedback conditions.

After the four weeks of treatment, a fifteen-minute recitation lesson was videotaped for each of the twelve interns. Class settings and pupils were constant across the pretreatment videotape sessions, the four treatment sessions, and the posttreatment videotape sessions.

Coding Techniques and Procedures

The twenty-four pre- and posttreatment videotapes were first coded for quantity of pupil versus teacher talk. Percentages of pupil and teacher talk were calculated for each videotape. The remaining six dependent variables (three levels of teacher soliciting and three levels of teacher reaction) were coded in two separate manners. The distinction between these two methods of coding was in terms of the unit of analysis assigned to the dependent variables. The first method coded *structural* teacher acts. If a question was formally conceptual, it was coded as conceptual whatever the level of the subsequent pupil response. The second method coded *functional* teacher acts by extending the unit of analysis to include the formal teacher acts plus subsequent pupil behavior. In the functional paradigm, a question which was formally conceptual was coded as such only if the subsequent pupil response was at the factual level, the structurally conceptual question was recorded as a factual question. Similarly, if a teacher reaction was formally and structurally a closing reaction, but the same pupil continued to talk; it was functionally coded as a sustaining or extending reaction depending upon the level of the pupil's subsequent vocalization. Percentages of factual, conceptual and theoretical questions; and of closing, sustaining and extending questions were calculated for both structural and functional coding methods.

All the coding was performed by the same observer-recorder without any knowledge of treatment assignments. A second observer coded one-third of the tapes using both the structural and functional coding techniques. Inter-rater agreements for each of the dependent measures were determined using the whole-session method (Repp, Deitz, Boles, Deitz, & Repp, 1976) averaged across the number of videotapes which both observers coded.

Results

Table 2 reports agreements for each of the dependent variables and coding methods. Teacher and pupil talk were not coded functionally as "amount of talk" is a structural variable by definition.

Two-way ANOVAs on the dependent variables pretreatment values

TABLE 2  
AVERAGE INTER-RATER AGREEMENTS FOR ALL DEPENDENT VARIABLES

Dependent Variable	Structural (%)	Functional (%)
Teacher Talk	89	--
Pupil Talk	87	--
Factual Questions	72	83
Conceptual Questions	77	75
Theoretical Questions	82	78
Closing Reactions	69	92
Sustaining Reactions	75	90
Extending Reactions	88	89



yielded no significant differences between treatment groups, for either functional or structural data, confirming that the random assignment procedures produced equivalent groups with regard to the variables under investigation. Two-way ANOVAs were then performed on the posttreatment values. One series of ANOVAs used the structural data, another series used the functional data. In the absence of pretreatment differences between groups, these analyses were sufficient to demonstrate the hypothesized differential outcomes attributable to the different coding techniques. Results significant at  $\alpha = .10$  are reported since this level of significance best represents the trends in the data.

Marginally significant *F* values appeared for the interaction effect between videotape feedback and feedback from structured observation instruments on the student talk and teacher talk variables. The treatment groups which received videotape feedback (V) or structured observation feedback (O) alone were associated with a greater amount of student talk (and therefore a lesser amount of teacher talk) than were the groups which used both (VO) or neither (C) feedback modalities.

TABLE 3  
MEANS FOR ALL TREATMENT GROUPS ON ALL POSTTREATMENT  
DEPENDENT VARIABLES

Variable	C Group	V Group	O Group	VO Group
Teacher Talk	72.20	53.77	57.13	70.07
Student Talk	27.80	46.23	42.87	29.87
Factual Questions (Structural)	76.67	63.67	31.50	61.67
Factual Questions (Functional)	62.33	65.73	62.37	78.90
Conceptual Questions (Structural)	14.33	30.67	45.83	20.00
Conceptual Questions (Functional)	35.60	24.40	27.93	20.03
Theoretical Questions (Structural)	9.00	5.67	22.67	18.33
Theoretical Questions (Functional)	2.37	9.86	9.70	1.07
Closing Reactions (Structural)	45.73	29.20	27.23	35.30
Closing Reactions (Functional)	73.20	85.30	86.27	71.53
Sustaining Reactions (Structural)	35.16	42.50	54.00	39.33
Sustaining Reactions (Functional)	19.00	10.81	10.53	19.47
Extending Reactions (Structural)	19.11	28.30	18.77	24.37
Extending Reactions (Functional)	7.80	3.89	3.20	9.00

Analyses of the structural data for teacher questioning and reacting variables yielded statistically significant results for factual questions and closing reactions. For factual questions, there was a significant main effect for the structured observation factor and a significant interaction effect between videotape and structured observation factors. Both results are attributable to a very low mean percentage of factual questions (as compared to the other treatment groups) associated with the structured observation alone (O) treatment.

The significant interaction effect between videotape and structured



TABLE 4  
ANOVA SUMMARY TABLES OF SIGNIFICANT RESULTS

	Source	Sum of Squares	df	Mean Squares	F ratio
Teacher Talk	Video (V)	22.688	1	22.688	.12
	Structured Observation (O)	1.141	1	1.141	.01
	Interaction (I)	737.900	1	737.900	3.88*
	Error (E)	1521.279	8	190.160	
Student Talk	V	22.140	1	22.140	.12
	O	1.268	1	1.268	.01
	I	741.040	1	741.040	3.88*
	E	1527.558	8	190.945	
Factual Questions (Structural)	V	867.000	1	867.000	2.50
	O	1865.720	1	1865.720	5.36**
	I	1489.791	1	1489.791	4.28*
	E	2784.656	8	348.082	
Closing Reactions (Structural)	V	53.763	1	53.763	.45
	O	115.320	1	115.320	.96
	I	453.870	1	453.870	3.78*
	E	960.952	8	120.119	

\*  $p < .10$

\*\*  $p < .05$

observation factors for closing reactions is attributable to a low percentage of closing reactions in Group O. This score was significantly lower than that of Group C which utilized neither feedback mechanism.

On the basis of the analyses performed on the structural data it would appear that the treatment group which received feedback from structured observation instruments (Group O) moved in the desired treatment directions more than any of the other three groups. Unfortunately, similar analyses performed on the data derived from functional coding did not isolate any significant treatment effects with respect to teacher questioning and reacting variables. It is this difference in outcomes as a consequence of different recording methods which forms the focus of the following discussion.

### *Discussion*

While the success of teacher training research as compared to teacher effectiveness research is no doubt attributable to several factors, it seems not unlikely that the structural-functional dichotomy illustrated herein plays at least some part. Structural teacher acts are most often the dependent variables in teacher training research. These variables have been successfully affected (in terms of changes in their emission frequencies) by a variety of training procedures and programs (cf. Gage & Berliner, 1975). The problem

has been that when these same structural teacher behaviors act as independent variables in teacher effectiveness research, they generally produce little systematic alteration in pupil learning outcomes.

The functional short term effects upon pupil performance of structurally-defined teacher behaviors have unfortunately tended to be ignored. While behaviorally-oriented educational psychologists (Becker, Englemann, & Thomas, 1971; O'Leary & O'Leary, 1972; Thoreson, 1972) have long advocated the notion of teaching as a series of tri-member contingencies in which the teaching act is functionally effective and meaningful because of its relationship to immediately antecedent and consequent pupil behaviors, more traditional research on teaching has not adopted a research strategy compatible with this conceptualization. If the teacher effectiveness researcher wishes to link teacher activity with long-term learning outcomes, he might be well advised to very carefully consider short-term interactive changes in pupil behavior as a function of discrete teacher acts. As mentioned earlier in this paper, it seems almost metaphysical to suppose that desirable long-term changes in pupil performance will materialize in the absence of such short-term alterations. If the whole is at least the sum of its parts, teacher effectiveness research should pay more attention to the "parts".

Berliner (1976) draws attention to this same problem in his discussion of problems associated with the determination of the "correct" unit of analysis assigned to the independent variable in teacher effectiveness research. Is the appropriate unit the teacher behavior, teacher behavior plus wait-time, teacher behavior plus wait-time plus student response, or something even larger? The importance of this question cannot be overemphasized if teacher effectiveness research is to be of any particular value in accruing knowledge about the nature of teaching and its effects on pupils. The training study included here resulted in changes in the emission frequencies of discrete teacher behaviors, but when these behaviors were functionally linked to actual consequent student acts, changes evaporated. Further, the a priori assumption that higher levels of teacher questions and reactions result in lower quantities of teacher talk (greater quantities of pupil talk) seemed to be easily inferable from the structural codings; but when the functional analyses came up empty, the impression was that dependent variables other than those studied determined the actual changes in talk ratios. Thus not only does the choice of different units of analysis result in different outcomes, it also leads to the formulation of inferences which may be more or less the product of a priori theorizing than of empirical reality.

Perhaps one of the most interesting suggestions for further study which emanates from the present consideration of methodology is that teacher effectiveness might be associated with the extent to which a teacher uses specific teaching skills which have the desired relationship to subsequent student behaviors. The ratio of "successful skill emissions" (in terms of student response) over "unsuccessful skill emissions" (again, determined by student response) might be found to be of considerable utility as an independent variable in its own right.

The implications of such ideas for teacher training might be that interns should be more carefully trained in discriminating specific classroom



## *Differential Outcome Effects in Teacher Training*

contexts in which particular skills are likely to have desired short-term effects and in assessing, on the basis of subsequent student activity, whether or not such desirable relationships actually do occur. At the very least, one can certainly posit that such discrimination and assessment skills are as important as the ability to emit a particular skilled behavior at a high frequency without sufficient contextual concern.

This project was supported by Simon Fraser University President's Grant 716-027.

### *References*

- Acheson, K. A. The effects of feedback from television recording and three types of supervisory treatment on selected teacher behavior. Unpublished Ph.D. dissertation, Stanford University, California, 1964.
- Becker, W. C., Engelman, S., & Thomas, D. R. *Teaching: A course in applied psychology*. Chicago: Science Research Associates Inc., 1971.
- Berliner, D. C. Microteaching and the technical skills approach to teacher training. Stanford, California: Stanford Center for Research and Development in Teaching. Technical Report No. 8, 1969.
- Berliner, D. C. Impediments to the study of teacher effectiveness. *Journal of Teacher Education*, 1976, 27, 5-13.
- Borg, W. R., Kelley, M. L., Langer, P., & Gail, M. *The minicourse: A microteaching approach to teacher education*. Beverly Hills, California: Macmillan Educational Services, 1970.
- Dunkin, M. J., & Biddle, B. J. *The study of teaching*. New York: Holt, Rinehart and Winston, 1974.
- Flanders, N. A. *Analyzing teacher behavior*. Reading, Mass.: Addison-Wesley, 1970.
- Gage, N. L. An analytical approach to research on instructional methods. In A. Yee (Ed.), *Social interaction in educational settings*. Englewood Cliffs, N. J.: Prentice-Hall, 1971.
- Gage, N. L., & Berliner, D. C. *Educational psychology*. Chicago: Rand McNally, 1975.
- Houston, W. R., et al. (Eds.) *Resources for performance-based education*. Albany, N. Y.: State University of New York, State Education Department, 1973.
- MacDonald, F. J. A theoretical model for the use of observational learning in acquiring teaching skills. Paper presented at the meeting of the American Educational Research Association, Los Angeles, February, 1969.
- O'Leary, K. D., & O'Leary, S. G. *Classroom management: The successful use of behavior modification*. New York: Pergamon, 1972.
- Perry, R. P., Leventhal, L., & Breen, L. J. The role of observational learning in the acquisition of effective teaching behaviors. Unpublished manuscript, University of Manitoba, 1975.
- Program on Teaching Effectiveness. Teacher training products: The state of the field. Stanford, California: Stanford Center for Research and Development in Teaching. Research and Development Memorandum No. 110, 1974.
- Repp, A. C., Deitz, D. E., Boles, S. M., Deitz, S. M., & Repp, C. F. Differences among common methods for calculating interobserver agreement. *Journal of Applied Behavior Analysis*, 1976, 9, 109-113.
- Rosenshine, B. *Teaching behaviors and student achievement*. New York: Humanities Press, 1971.
- Simon, A., & Boyer, E. G. (Eds.) *Mirrors for behavior: An anthology of classroom observation instruments continued*. Philadelphia: Research for Better Schools, Inc., 1970.
- Thoreson, C. E. (Ed.) *Behavior modification in education*. Yearbook for the National Society for the Study of Education, 1972.
- Winne, P. H. The assumption of equal units of effects in research on teacher effectiveness. Unpublished manuscript, Simon Fraser University, 1976.



A. G. KONRAD

*The University of Alberta*

## College Trustees View Their Selection

*Community college trustees in Alberta are appointed by the provincial government. Very little is known about the selection procedure; the purpose of this article is to examine the present methods of selection and how trustees view them.*

*Research is reported which ascertains the perceptions of 35 newly-appointed board members. Most of the trustee respondents expressed strong satisfaction with the present selection procedures and with the provision for faculty and student memberships. However, there was some dissatisfaction with the manner in which the chairperson was selected, and the fact that college presidents have voting memberships on their college boards.*

*Concerns were noted by the respondents regarding lack of information about trustee selection procedures, dangers associated with political patronage, time delays in making appointments, presidential memberships on boards, "rubber stamping" faculty and student appointments, and the method of selection of board chairpersons. Suggestions were also made about possible means of improvement. These suggestions dealt primarily with means of removing the political overtones from the process and providing better information about selection procedures and criteria. (Dr. Konrad is Professor of Educational Administration at the University of Alberta.)*

Members of postsecondary governing boards in Alberta are appointed by the Lieutenant-Governor by an Order in Council upon the recommendation of the Minister of Advanced Education. How newly-appointed college board members view their selection and the overall process was the topic of a recent research project at the University of Alberta (Konrad, 1974).<sup>1</sup> In this article the present procedures of selecting college board members are briefly reviewed following which the views of newly-appointed board members regarding these procedures are examined.

### *Alberta College Trustees*

Public postsecondary institutions in Alberta are provincial in scope; they do not serve geographically defined regions. Although governance patterns have changed somewhat throughout the development of the current system

and vary also among the different types of institutions, today all postsecondary institutions are responsible to government through the Department of Advanced Education.

The six public colleges established prior to 1975 operate under the direction of governing boards comprised of five lay members appointed by government, the president of the college, one faculty and one student representative. Public members serve three-year terms, once renewable, and are appointed on a rotational basis. Institutional members serve one-year terms which may also be renewed. The chairperson of a board also serves a three-year term by government appointment. All terms begin on July 1.

### *Research Procedures*

The purpose of the study was to ascertain the views of newly-appointed board members about their selection and the overall process. No attempts were made to examine the role and functioning of college boards; this study was delimited to an examination of the process followed in selecting new board members.

During the eighteen-month period of the study, ending in the spring of 1974, thirty-seven persons were newly-appointed or re-appointed to college boards in Alberta. All but two of the trustees agreed to cooperate in the study. Of the 35 board members who participated, nineteen were lay members appointed by government from the public-at-large, eight were faculty and eight were student representatives. Because college presidents serve on their boards by virtue of office and, consequently, cannot reflect on the question under study through personal experience, they were excluded from the study. Trustees who had not been appointed recently were also excluded, because their perceptions may have lost some precision over time.

Board members were interviewed in person to determine their perception of the procedures used in their nomination and selection, to solicit their opinions regarding institutional membership on boards and to obtain suggestions for improving selection procedures. All interviews were tape recorded and full transcripts prepared for content analysis.<sup>2</sup>

### *Research Findings*

Eight general questions were used as a guide in interviewing. The first two focused upon background experiences and were somewhat preliminary; the remaining six centered on the personal perceptions of board members regarding their selection.

#### *Background Experiences*

1. *What community activities were you engaged in prior to your appointment to the board?* All but one of the 19 lay members had engaged extensively in some kind of community activities. Seven had been members of service clubs; three had served on school boards; and others had participated in groups such as civic advisory boards, chambers of commerce, business and professional groups, parent-teacher associations and church organizations.

In addition to their involvement in community activities on a voluntary basis, nine respondents (47%) had held public office at the municipal level

and three others had held positions at the provincial level. Of these, 50 percent were elected offices. Board members appear to be drawn from the "active" sector of society, providing public service both through formal and informal avenues.

But board members do not present a uniform background. One lay respondent offered the following:

I really think that the Lieutenant Governor in Council, perhaps with the advice of . . . , made an effort to find people . . . with different backgrounds and with different interests and from different religious and ethnic backgrounds.

Previous involvement in some aspect of community life prior to board appointment appears to characterize newly-appointed college trustees.

2. *What was the nature of your involvement in party politics?* Table 1 indicates the political party associations of newly-appointed board members. Although none had held elected governmental office, a majority (63%) of those interviewed had served in some official capacity within the various political parties. Two had been candidates for political office and three had served in executive positions. Only seven indicated no previous affiliation with party politics.

TABLE 1  
POLITICAL INVOLVEMENT OF PUBLIC MEMBERS (N = 19)

Nature of Involvement	No.	Response	
			%
Candidate for office	2		10.5
Executive officer	3		15.8
Campaign manager	4		21
Fund raiser	1		5.3
Party member	2		10.5
No affiliation	7		37

Views regarding the significance of party politics in the appointment of public members varied considerably. One trustee who had been charged by faculty that her appointment was a "political one" acknowledged, "I can understand why people would think this, with the association that my husband and I have with political figures." Another confided, "I just happen to be a woman from the right political party with the right experience and, presumably, they wanted some type of balance, so they picked me." She hastened to add, "If it was just political, we would have an awful problem." Political patronage appears to be present in the appointment of college trustees, but the real concern lies with whether or not it results in the appointment of poorly qualified persons. A rather caustic criticism was levied at the government by a faculty board member:

I think it's ghastly, you know. A guy gets stuck on something, because he ran as a political candidate . . . and didn't make it. And he hasn't got much of a job, and they stick him on this thing as a holding action.

Similarly, a lay member emphasized the necessity of caution in making "political" appointments:



It depends so much, I think, on the integrity and the ability of the Minister or those people ultimately responsible for the appointments, as to whether or not people who are appointed as political hacks are just given an honorary position because of past political service.

While the dangers of political patronage were readily acknowledged, several respondents disavowed its influence in recent appointments.

Well, it's always been my belief that political affiliation should have nothing to do with government appointments . . . , particularly so with respect to an educational institution. I'm personally satisfied that it didn't have any influence with regards to the appointments at [this] college. (Public member)

A student member, reflecting upon her board experiences, replied,

I think we're very fortunate in that our board does not seem to be a particularly political kind of group. . . . I got the impression that in other places [of Canada] it's much more strongly a political kind of appointment.

So much for political involvement, its influence and potentially detrimental effects on the selection of board members.

Personal Perceptions

3. *Can you comment on your nomination and appointment?* Vacancies on college boards are announced in the major newspapers of Alberta when nominations are being received by the government. All but two of the public members were aware of the person or group that had nominated them for office. As Table 2 shows, seven (37%) indicated that they were nominated by a Member of the Legislative Assembly (MLA). In the majority of these cases, the contact had been made by a local MLA, but at least in three instances the Minister had become personally involved in the nomination. "I think that I was drafted, pure and simple," responded one board member. "He [The Minister] put the pressure on that he'd like me to consider it." But pressure came not only from government sources. One newly-appointed board member recalled that he was "kind of forced into it" by another board member.

Information regarding actual selection procedures is nonexistent. Respondents knew neither the steps that had been followed nor the criteria that had been used in making the final selection from among the nominees.

TABLE 2  
NOMINATION SOURCE OF PUBLIC MEMBERS (N = 19)

Source of Nomination	Response	
	No.	%
Member of Legislative Assembly	7	36.8
Service Club	3	15.8
College board member	2	10.5
Professional colleague	2	10.5
Other (Citizens' committee, school board, self)	3	15.8
Unknown	2	10.5

(It may be assumed that persons nominated directly by a government source felt assured of their appointment.) As one respondent put it, "I think some of the mystery of it could come out of it [the process]." No dossiers had been required and no interviews conducted. There was "no communication from anybody" between the submission of the nomination and the contact to ascertain the willingness for service.

Generally, successful candidates were contacted by telephone and asked whether they would accept the appointment if selected. Notification of appointment itself usually followed at short notice. From these interviews it appears that the government is centrally involved in the selection procedures — even to the extent of keeping to itself information about the process and the specific criteria employed.

4. *How do you feel about the selection process involving public members?* Table 3 reveals that 79 percent of the public members expressed general

TABLE 3  
OPINIONS ON SELECTION PROCEDURES

Response Category	Public Members N=19	Faculty N=8	Students N=8
Satisfied with public selection	79%	50%	12.5%
Dissatisfied with public selection	21	50	75
No response <sup>a</sup>	0	0	12.5
Satisfied with institutional selection	63	100	100
Dissatisfied with institutional selection	5	0	0
No response <sup>a</sup>	32	0	0
Satisfied with number of institutional representatives	52	25	50
Increase faculty representation <sup>b</sup>	21	37.5	0
Increase student representation <sup>b</sup>	21	50	25
No response <sup>a</sup>	26	12.5	25
Favour presidential membership	11	25	12.5
Opposed to presidential membership	26	25	12.5
No response <sup>a</sup>	63	50	75
Satisfied with chairperson selection	16	0	0
Dissatisfied with chairperson selection	58	0	0
No response	26	100	100

<sup>a</sup> In most instances this category includes respondents who did not comment on this item--the interviews were only partially structured.

<sup>b</sup> These are not mutually exclusive categories; some respondents were counted twice.

satisfaction with their selection; faculty opinion was equally divided on the appropriateness of the selection procedures, while students declared strong dissatisfaction with them. The advertising of appointments, so that any citizen may put forward the name of someone who he thinks is suitable, in the opinion of a public member, is the most attractive feature of the process.

It was readily acknowledged that the selection procedures placed a lot of power into the hands of the Minister. That appointments are largely political was acknowledged by 36 percent of the public members. "He [The Minister] is bound to promote his friends," criticized a student. A public member countered,

But it's a hard thing to stay away from because, really, a politician is going to appoint someone he knows. He's going to appoint somebody he thinks is capable. Well, let's face it, the people that he's going to know best are the people he's worked with.

The most obvious alternative to the appointive process would be an elective procedure. Election of board members was favoured by 18 percent. Generally, support for adopting an elective approach was based upon the view that it would be "a more democratic" way of selecting trustees. However, if the process were to be elective, countered one public member, "I think it would destroy a great deal of its effectiveness."

Three major factors identified in support of retaining the appointive process related to constituency, cost and the quality of the candidate. Since community colleges are provincial institutions, it would be difficult to identify the constituents for a general election, suggested a public member. Furthermore, it would be very expensive to conduct election campaigns, and many trustees felt that such costs would not be warranted.

The primary objection to selecting trustees by popular election had regard to the quality of the candidate. Indeed, a number of trustees admitted that they would not have submitted themselves to the "scarey," "difficult" hassles of campaigning. "Quite often," summarized one respondent, "people that would make excellent board members would never seek an election."

Some also felt that appointed members had stronger service motives than elected members would have. "Now I consider it more of a service," testified a respondent, "whereas if I had run for the position then I would think it could be more of a personal ambition." Of course many disagreed with these views. A faculty member rejoined, "The thing that bothers me is that the people who are appointed to the board seem to take it as an honour rather than as a duty."

Two complaints regarding the present procedures were frequently expressed — one had to do with the selection criteria and the other with timing. "My only comment is that I don't know very much about it," responded a faculty member. And then he explained, "I don't know what sort of criteria are used for the selection of public members, nor do I know how the criteria are applied. I'm not even sure who makes the final selection." Similarly, a public member thought it would be a good idea "for the other board members to know how people came on, and even for the community at large" to know more about the process.

With respect to timing, respondents charged government with inordinate delays in making appointments. Evidently this was true not only of public



members, but also with respect to the “formal” appointment of institutional members. A public member recalled an instance where a student had no vote “because there had been no letter to the board of her appointment.” Somewhat sarcastically the member continued, “Why do they have to rubber stamp it in Edmonton?” In general, however, the present procedures were regarded favourably by newly-appointed trustees.

5. *What are your views regarding faculty and student membership, and the procedures involved in their selection?* Almost without exception, respondents commented favourably on faculty and student membership on governing boards. The following comments are representative of enthusiastic support for institutional memberships:

I think it is imperative that you have student representation. Let’s face it, without students there wouldn’t be a college, and there wouldn’t be teachers. (Student member)

The representatives that we’ve had from both faculty and the students have been excellent. They were very good, very conscientious and, certainly, they were able to bring points of view to the board that would otherwise not have been there. (Public member)

In the discussion of institutional memberships, respondents generally agreed that faculty and students served as full-fledged board members. As one public appointee put it, “A board member is a board member, whether he is a student member or a faculty member or a community member.” Faculty members were particularly persuasive on this point. “I think that it is not the role of the person to serve as a representative of the Faculty Association,” maintained a faculty board member, “but more as a faculty person who sits on the board.” Faculty and student members provide unique perspectives to board discussions, but they do not regard themselves “accountable” to their respective associations. “What I vote on in a board meeting may not necessarily be good for the students in one particular instance,” conjectured a student member, “but it is good for the whole community.”

Some factors apparently make it difficult for institutional members to function effectively. Occasionally, issues arise which appear to place institutional members into positions of conflict of interest. An example of such a situation was seen in the negotiation of faculty salaries.

Another impediment to effective institutional membership on boards identified by respondents lies in the length of faculty and student terms. A faculty member theorized,

Faculty should be appointed for a longer term of office. They could serve a three-year term just like the public members do. In this way their effectiveness would be greatly improved. By the time faculty and students are oriented to the job, the year is almost over. It also takes time to establish one’s credibility.

But if this is a matter of concern for faculty where provisions exist for the renewal of appointments, how much more critical must this be for students. In most instances, boards lose their student members during the summer months. But, again, respondents favoured institutional memberships, even with their shortcomings.

About half of the public and student members were satisfied with the number of institutional members; only one-fourth of the faculty expressed

similar satisfaction (Table 3). Some public and faculty members felt that the number of faculty members should be increased, and some respondents from each group favoured an increase in the number of student members. Reasons offered for increased institutional representation related to concerns for parity and effectiveness. The latter was raised particularly with respect to increased student representation. A public member surmised that a student might not have the level of personal confidence of other board members, and so it might be wise to have two student representatives "so that they could provide moral support for each other." Perhaps how effectively a student member serves on a board depends not only on the student member, but also on the other members of the board and their willingness to work together for the common good of the institution.

All of the faculty and students and a majority of the public members interviewed were satisfied with the procedures used in selecting institutional members to college boards (Table 3). Public members were largely content to allow faculty and students to "pick their own" representatives.

In each of the six colleges, the faculty elected their representative to the board of trustees. Procedures varied from formal written nomination papers, with a mover and a seconder, posted about a week before ballots were cast to a nomination from the floor without a selection committee and a secret ballot that took place immediately. Although procedures varied, the main concern in conducting faculty elections was that "the faculty certainly has a chance to select whomever they wish."

Procedures for selecting student representatives for board membership during the time of this study were of three types: ex officio, by appointment and by election. Two student body presidents served as members of their college boards. The advantages in this arrangement, according to these respondents, lay in an increased measure of their effectiveness as campus spokesmen. "I'm more well informed about what is happening. I get a clearer picture on how each part relates to each other," offered one student body president.

Student representatives were appointed by student council at two colleges, and elected by the total student body at the other two colleges. Opinions differed regarding which was the best way to select a student representative. "In my opinion," asserted a student member, "an election is essential." Another added, "The election system at least identifies in the minds of some of the students who is the Board of Governors' representative." In opposition to elections, an appointed member propounded, "In the big colleges . . . you get about three kids running for a position, and most of the kids don't know whom they're voting for." And, commonly, only a fraction of the students were reported to participate in elections. Of course, for some, board membership was viewed as a political process, and these appreciated the opportunity "to go through the electoral system and sweat it out."

The Alberta experience with institutional membership on college boards has been a good one, at least from the perspective of new board members. While selection procedures differ somewhat among the institutions, faculty and students seem to regard their memberships seriously and are generally intent on finding the best possible representatives to serve on boards.



6. *What opinions do you hold of the college president's membership on the board?* By legislation, Alberta college presidents serve as voting members on their governing boards. Opinions on this matter were solicited only during the latter part of the study and, consequently, the findings may be less definitive. At best they may identify concerns that deserve further exploration.

Only a few respondents favoured presidential membership on college boards (Table 3). A faculty member felt this arrangement forced the president to clarify his position on agenda items. Another faculty member felt that the arrangement seemed to be working "fairly well," but added, "it may be different with a different president." Even these favourable opinions were not very positive in their support of voting membership for the college president.

Five of the seven public members who commented on this matter were opposed to presidential membership on governing boards. Generally, opposition was on a "theoretical" rather than a "practical" basis. And, clearly, it was not regarded as "a personal thing." It was questioned whether or not it was "healthy" for the president to actually have a vote with respect to his own decisions and actions. "If he is a true agent of the board," queried a student member, "then why is he allowed a vote on the board?" This position was further propounded by a faculty respondent:

As the president, he should be the person who implements and executes the policy decisions of the board and, to that end, he should be present as a resource person, and keep the board informed on a number of things that they may be doing contrary to the Act or past policies. But he should not be a voting member because of the repercussions that this tends to have.

In simple words, a number of respondents saw a "conflict of interest" in the president's board membership.

No one questioned the importance of the president's participation in board meetings. There was considerable concern, however, about the nature of the president's participation. Indeed, one public respondent expressed alarm over the president's ability to "sway a board." When that occurs, she complained, "I feel that what happens to you is that you become nothing more or less than a rubber stamp."

Concerns over this issue should challenge presidents to develop adept skills in exercising their board membership responsibilities. Indeed, opinions regarding presidential membership on governing boards call into question the basis for such a provision.

7. *How do you feel about the selection of the chairperson of the board?* Under present legislation, a board chairperson serves a three-year term by government appointment. Only public members were asked to express their feelings about the selection procedures involved. All but three of the persons who responded to this question, as shown in Table 3, expressed dissatisfaction with present arrangements.

One argument for retaining the present procedures was advanced in the interest of maintaining good relations with government. "If a board appointed its own chairman," surmised a respondent, "you might get some friction between the chairman here and the government, and I think that would be undesirable."



The majority of public members felt that the board should be either consulted prior to the selection or given sole responsibility for making the selection. Some thought that board members should be asked to "suggest somebody," or at least help the Minister in "screening" candidates for the position.

A number of board members argued that it would be "more democratic to have the board choose its own chairman." Furthermore, offered a public member, "We are the people that are going to know whether that guy is going to be able to handle the chairmanship or not." And finally, another confided, "If the people you're working with think enough of you to elect you chairman, then usually you are the best person available in that group." Presumably, no one would be appointed as chairperson without previous experience as a board member.

8. *In what ways might the selection of board members be improved?* Although many answers to previous questions contained implications for the improvement of selection procedures, respondents were given a specific opportunity to suggest changes that could improve the process. Whether proposed changes would actually improve the quality and effectiveness of boards must, of course, remain a matter of conjecture. The wide-ranging series of suggestions offered by respondents will be summarized in four categories related to matters of information, involvement, instrumentality, and institutional membership.

*Information.* An over-riding concern of respondents dealt with the general lack of information regarding the selection process. Most people were evidently quite ignorant about selection procedures. How persons are selected for board membership should be clarified.

Information should provide a general understanding of the procedures, but it also should set forth specific criteria for the screening and selection of board members. A faculty member proposed,

These criteria should be arrived at as a result of a careful examination as to what is going on at the college, [and] what the functions of the board are. Then various individuals should be judged against these criteria.

Making the criteria public would eliminate the nomination of persons who were obviously not qualified for board membership.

Board functioning also could be improved through a more intense orientation of newly-appointed board members. Some board members, criticized a faculty member, "are still fumbling as to the function of the board" after they've served for a year or two. And a systematic orientation program could be used to provide them with basic information about the college and the role of its governing board.

*Involvement.* A second set of suggestions related to the involvement of a wider spectrum of persons in the selection process. "Who makes the selection?" concerns itself not so much with legal provisions, but perhaps far more precisely with how the process functions. Many public members expressed frustration with their lack of involvement in the selection process. The nature of the proposed involvement ranged all the way from getting out into the community to "look for people who would be able to do a job in the education field for the college in this community" to interviewing and "screening" candidates for recommendation and for appointment. One

public member advocated involvement on the basis of the knowledgeable ability of board members:

We do know the people here in the community, and we do know the present board and the problems that are on the board, and what type of people you would like to see coming in. As certain people leave there are going to be certain gaps. . . . If certain names did go forward, we might say that one person might be more valuable on the board than another, from our point of view.

A faculty member, however, warned against the danger of board involvement: "They might be inclined to appoint a friend or an acquaintance who would only tend to perpetuate the same kind of thing. This could be a mistake."

Another proposal for greater involvement suggested the creation of a nonpolitical selection committee. A faculty member felt that a "formal selection committee" could "solicit the willingness of people to run." "Considerable time and effort should be spent in screening and selecting suitable candidates," advised a public member. Indeed, a faculty member suggested that at least three candidates should be interviewed for each position that was to be filled. Furthermore, another faculty member argued,

There must be people within the community with a certain adeptness in screening a large number of candidates. Certainly that could be done far more easily locally than by somebody that's in Edmonton.

The major concern in all of these suggestions appeared to be with finding a more effective way of identifying the best possible persons for board membership and, secondarily, to make the process at least appear to be more impartial by removing it from the "personal and political whims" of the Minister.

*Instrumentality.* Thirdly, respondents had suggestions regarding the instrumentality or logistics of the selection process. Perhaps the procedures were quite acceptable, but they just didn't seem to be functioning too well. Repeatedly, complaints were received about the long delays in making appointments. "If it's the system that creates this delay, then it's got to be faulted for that," judged a public appointee. In plain words, respondents viewed the process as having been "bogged down," and efforts should be made to announce new appointments well in advance of their effective dates.

Perhaps the resolution of the time-lag problem for institutional members comes from another proposal of a public member. Student and faculty representatives could serve "automatically" upon their selection. Why must their selection be "rubber-stamped" in Edmonton? Furthermore, it was proposed that students select their representatives in spring, perhaps April, so that they could participate more easily in board sessions during the summer.

Another functional matter related to the organization of the board. The majority of public members expressed dissatisfaction with present procedures, and recommended direct involvement in the selection of the board chairperson. It was also recommended that the board itself should establish policy regarding the term of office for the chairperson. Another public member regarding the present three-year term "a bit limiting" and



advocated the adoption of a rotational system for chairpersons within the board.

*Institutional memberships.* Finally, several respondents felt that changes with respect to institutional memberships could improve college boards. As noted earlier, a majority of public members and about half of the faculty and students who spoke to the issue of presidential memberships were opposed to it. Further consideration should be given to the effect of this provision on board functioning.

Members of all three constituent groups called attention to the number of institutional members on governing boards. It was suggested that an additional student member on the board, and perhaps also an additional faculty member, would increase the effectiveness of institutional members. The most "liberal" proposal came from a student who recommended "five students, five public members, and five faculty" constituting the board, and the president and his staff serving as resource people. A faculty recommended a "2-to-1" relationship of public members to institutional members. "Then, of course, when you have four lay members, you should have two faculty and two students."

The length of faculty and student terms also was raised for discussion. In the interest of providing continuity, a faculty member urged that the term be extended to at least two years. One faculty respondent argued for the retention of the one-year term on a renewable basis because, he maintained, this was a way of giving "a vote of confidence to the guy who was doing a good job." Others felt that the orientation to the position in itself warranted a longer term of office.

### *Conclusions*

This article reviewed the procedures for selecting college board members in Alberta, and reported the perceptions of newly-appointed board members of the nomination and appointment procedures followed. Although the study was limited in scope, a number of major conclusions appear to be noteworthy.

1. College board members are drawn from an active sector of society, representing perhaps a somewhat limited cross-section of their communities.

2. Involvement in party politics prior to selection for trusteeship is common among public members.

3. Members are nominated by a variety of sources, although government involvement in the process seems to be quite substantial.

4. The selection procedures, and particularly the criteria followed, are unknown to the majority of newly-appointed trustees. Nevertheless, most respondents expressed satisfaction with present procedures used in the selection of public board members.

5. The provision for faculty and student membership on governing boards in Alberta is regarded very favourably by all respondents. If changes are to be considered, it should be in the direction of increasing the number of institutional members.

6. Most respondents were opposed to the provision of voting membership for college presidents on their boards.



7. Most public members expressed a desire to be directly involved in the selection of the board chairperson.

8. Major suggestions for improving the process were received in areas relating to the dissemination of specific information about selection procedures and criteria; the involvement of a wider spectrum of persons in the nomination and appointment procedures, particularly in an effort to remove the process from the personal and political whim of the Minister; increased efforts to speed up the instrumentality or functioning of the process; and a reallocation of institutional representation on boards.

Trustees devote a great deal of time and effort to their college responsibilities. The study reported in this article provides strong evidence of satisfaction with present selection procedures and offers some suggestions for improvement. Better selection procedures should result in a higher quality of board membership and, consequently, an improved environment for teaching and learning in Alberta's community colleges.

1. The writer gratefully acknowledges a grant from the University of Alberta Research Grants and Contracts Office to assist in the data collection and analysis in this project.
2. Bevis Peters, a doctoral candidate in the administration of postsecondary education, prepared the initial content analysis of the interviews.

#### *References*

- Konrad, A. G. Selection processes of Alberta college trustees, 1972-74. Unpublished research report. Edmonton: The University of Alberta, 1974.
- Konrad, A. G. Community college boards: A Canadian perspective. In R. C. Richardson, Jr. (Ed.), *New directions for community colleges: Reforming college governance*. San Francisco: Jossey-Bass, 1975.

J. McLEISH

University of Victoria

## Knowledge and Revolution: The Student Movement in Tsarist Society (1856-1881)

*The prototype of the student activist is found in the discussion circles which flourished in Tsarist Russia in the 19th century. These arose spontaneously in the universities of Moscow and St. Petersburg following the defeat of Napoleon (1812) and the execution of the Decembrists (1825). Not initially conspiratorial, these groups became so as a direct consequence of arbitrary police action.*

*The movement became a serious threat to the Tsarist System after the Crimean War. The revolutionary students organised to educate the illiterate peasants (about 80 percent of the population) with a view to inducting them into the social movement. Over two decades (1856-1873), the panacea of popular education was a leading motif of the movement for reform. The Sunday Schools and the movement "To the People!" were both inspired by the ideas of the London émigrés, Herzen and Ogaryov. The ruthless suppression of these activities by the Tsar drove the students to terrorism, resulting in the assassination of Alexander II in 1881. (Dr. McLeish is Chairman of the Division of Psychological Foundations, Faculty of Education, University of Victoria.)*

The leading characteristic of Tsarist society, this "Kingdom of Darkness" (Dobrolyubov, 1859), was an almost total absence of human feeling. The barriers erected by social structure served to maintain individuals at a distance. The correct attitude to other people was a formality which concealed indifference. In a letter to a friend, Chekhov writes:

In this world it is essential to be indifferent. Only those who are indifferent are able to see things clearly, to be just and to work. Of course, I am only speaking of intelligent people of fine natures; the empty and selfish are indifferent enough anyway. (Quoted by Leites, 1953, p. 187)

The dangers of emotionality could be avoided by immersing oneself in paperwork. Official society under the Tsars lived by paper.

You sit in your office, no one confuses you, no one disturbs you. . . . A syllogistical

office activity gives me inexhaustible pleasure. One is drawn to it as to drinking or opium smoking. . . . If you think we have anything to do with these filthy scoundrels (the peasants) you are mistaken. There are secretaries and so on for that. (*ibid.*, p. 574)

The relations between official society and the universities during this period are well-conveyed in an anecdote of Professor Pavlov, the well-known liberal and the centre of progressive opinion in Kiev. In 1847, the University Chancellor, Governor-General Bibikov, declared:

Professors! You are free to meet amongst yourselves, but only to play cards. Students! Remember that I will turn an indulgent eye on dissipation. But a soldier's uniform is ready for anybody who indulges in freethinking. (Venturi, 1952, Vol. 1, p. 366)

Tsar Peter I introduced the sciences to Russia, especially those having to do with war. Initially supported by foreigners who had settled in Moscow, Peter was soon the centre of a wave of innovation. This was designed to further his ruthless drive towards imperialist expansion, described by foreign historians as the "Europeanisation process".

Peter's aims were imperialistic and militaristic: in pursuing them he was largely successful. But if Europeanisation means mass education (as in Prussia) or the development of peasant proprietorship (as in Germany and Switzerland) or the sponsoring of an independent, industrialised economy (as in England) Peter, and his successors, can be written off as totally ineffective. The limitations of his system of innovation "by Decree" can be illustrated from the field of education. The ukase of 1719 (Leary, 1919, p. 33) made literacy compulsory for the nobility, the clergy, and their children, this ability to be tested by examination. The illiterate were to lose certain privileges associated with their rank. Indeed, the final sanction was declassification and conscription into the army of life. Above a certain rank illiterates were forbidden to marry. But the problem of mass (peasant) illiteracy was not even raised by Peter—this fell into the area of private philanthropy. Autocracy willed the end of Europeanisation, but drew back from providing the means. Five years after the decree there were still only 110 elementary schools for a population of thirteen million (Chekhov, 1923, p. 22; Johnson, 1950).

Peter bequeathed the problem of controlling the process of change he had started to his successors. The question was: How to ensure that social and political changes are introduced gradually, by decree, "from above", and not sponsored "from below", in the manner of the rebel Pugachov? Reform as a government-sponsored monopoly, or revolution as an activity of the private entrepreneur—this was the question.

### *Tsarist Society in the Nineteenth Century*

#### *The Ruling Powers—The Conservative Ideology*

Tsarist society was meticulously graded according to social origin and, after Peter, bureaucratic function in the state service. The legal code set out in monumental detail the rights and duties of each grade. Class barriers were strengthened by minute definition of the obligations, the rights and mutual relations of each social class. There were about fifty grades in the civil service defined by function, by forms of address, by specifications of the



medals and distinctions open to incumbents at each grade. (*Bolshaya Sovetskaya Entsiklopediya*, 2nd ed., vol. 41, p. 447).

Women had very low status. For example, the husband had the right, given by an early lawgiver, to chastise his wife, provided he used a cane not thicker than the middle finger. Under Nicholas the First, barbarous physical punishments were meted out, especially to state criminals (Masaryk, 1919). Castigation with the knout, flogging in the Army and in prison, running the gauntlet, branding on the face, forced marches from Moscow or Saint Petersburg to Siberian exile—"a yawning abyss of all possible horrors" in the words of an official investigation by Baron Korff in 1844 (Wallace, 1912, p. 584)—these were common civil and military sanctions.

In the case of political offenders, public execution by hanging or shooting; solitary confinement lasting twenty, thirty or forty years; civil execution (a grotesque simulation of the real thing); hard labour in the mines; exile to Siberia—these were still common punishments.

In cases of treason the Tsar would personally interrogate the accused, he would prescribe the sentence. He behaved literally as "the Father of his people," especially when punishment had to be prescribed.

The ruling class was an authoritarian, father-identifying group. The pattern of punitiveness, authoritarianism, male chauvinism, emphasis on status differences and hereditary privilege—these persisted throughout the 19th century. In addition to fatherworship, the bureaucracy was devoted to militarism as a way of life. In performing their official function everyone wore military uniforms designed by the Tsar himself. Students and professors were not exempt. Some slept on camp beds in imitation of the Emperor.

The ideal of the Russian monarchy was summed up by Uvarov in 1837 in the slogan: *Autocracy, Orthodoxy and Nationalness*. This political, religious, and cultural ideal proved more and more inadequate as new needs were fostered by contacts with Western Europe. The State routines and value-system diverged more and more from the standards accepted by the educated classes. The ideas of the older generation were totally unacceptable to the young intelligentsia. The economic system proved more and more inadequate. A stagnant agricultural system could not meet the basic needs of an expanding population. A radical transformation was clearly necessary to provide the primary necessities of life for the masses of the people (Lyashchenko, 1956). Serfdom was an obsolescent social and economic system a century before it was abolished.

According to Lenin, *the production of grain for sale* was the signal for the collapse of the whole system of feudal agriculture (Lenin, 1957). The abolition of the Corn Laws in England in 1846 was the economic catalyst which accelerated the development of capitalism in industry and in agriculture. The need of the new system of production set in train vast social and economic changes in Russia. The same economic forces which opened up the American West enormously stimulated the growth of capitalist relations to the Russian towns and countryside.

An ideology of change, based on objective needs, confronted an ideology of conservatism, based on vested interests. The ideologists of conservatism became more and more incapable of governing in the old way. They proved

unable, and unwilling, to learn new ways. They become “superfluous men.”

At the appropriate moment they were discarded, leaving behind nothing but their illusions and woeful heritage of rigid bureaucracy.

*The Peasants before and after Emancipation (1861)—  
The Revolutionary Situation*

The peasant remained the main hope of the revolutionaries until the end of the nineteenth century. However, the peasantry showed a total lack of comprehension of the motives of the revolutionaries. It was an unwritten law that no peasant would ever appeal to the authorities against another peasant. But criticism of the “Little White Father” by the urban intellectual, trying to pass himself off as a peasant, was so disturbing that the normal outcome of the encounter (as in the movement “To the People!” of 1873) was police action, instigated by the peasantry. The peasants often took the law into their own hands and drove the educator-revolutionary from the village with blows (Meijer, 1955).

The folk tales (byliny) which centre around the mythical Ilya Muromyets have often been taken as a characterisation of the Russian muzhik (Ralston, 1872). They have as motif certain cyclical sequences which were said to characterize peasant behaviour. There is an alternation of jollity and melancholia, of absolute inaction and feverish bursts of activity, of fasting followed by feasting, long periods of repentant sobriety alternate with orgies of drunkenness, uncontrolled fits of physical violence and aggression are followed by tender care tinged with remorse, an insatiable curiosity is suddenly replaced by a doctrinaire censorship of expression, moods of passivity and fatalism supervene on a superabundant, indeed limitless, optimism about the capabilities of mankind, an untidiness and lack of system is punctuated by intermittent aspirations towards “kulturnost” (delicacy and fastidiousness) (Wallace, 1912).

It was exactly this unpredictability and resort to violence, coupled with their vast numbers, which persuaded the Russian revolutionary that the peasantry represented that primeval, ubiquitous force which alone could topple the Tsarist regime.

On 19 February 1861, the long-awaited manifesto liberating the serfs was published. The liberal opposition to the Tsar crumbled in face of this recognition of their demands. Other groups—the revolutionaries, the peasants, the student circles—did not take this view at all. When the provisions relating to compensation of the owners were clear, numerous disturbances in the shape of riots, jacqueries, murder of proprietors took place. There was a general administrative breakdown. In 1861 there were over 279 outbreaks, many on a scale which revived memories of Pugachov. For example, the Bezdna disturbance involved over five thousand peasants and two hundred and thirty soldiers. The outcome was fifty-one dead and seventy-seven wounded (Venturi, 1952, chapter 7). The commotion was due to the fact that the illiterate peasants accepted the fantasies of a semi-literate peasant as a translation of the Decree. The village priest, who was responsible for reading and explaining the Decree, was often semi-literate or even quite illiterate.

Following Emancipation, *obrok* (quit-rent) and *barshchina* (forced



labour) remained in force. Indeed, these became more burdensome. The landlords were compensated for the land which they sold to the peasants at inflated prices. In addition, the peasant had to pay an indemnity to his landlord for his freedom. Twenty million serfs emerged from Emancipation “free”, but indebted by mortgage for the rest of their lives (Lyashchenko, 1956, chapter 20).

Chernishevski predicted in 1857 that “nothing but abomination would result” from emancipation “from above”. He lived to see this prophecy realized. His *Letters without an Address* (a series of open letters on the results of Emancipation) were absolutely forbidden publication. Shortly afterwards he himself was arrested (7 July 1862) and convicted of criminal conspiracy.

The Tsar and his advisers were driven to Emancipation by the threat of revolution and by hard economic necessities. The classic statement of Marx about social change as the result of the productive *forces* maturing within the womb of the old society and being fettered by the productive *relations* (McLeish, 1968) is clearly exemplified in the social history of Russia. Emancipation was a prerequisite for industrial capitalism but the new system hardly improved the condition of the peasantry. In fact, they were worse off when feudalism was abolished. Many were crippled with redemption payments often five times higher than their real income (Venturi, 1952, Vol. 2, Chapter 7).

The Decree of 1861 was a compromise, in which everything was done to safeguard the landowners’ interest. They had wanted emancipation without land, only in that respect were their wishes disregarded.

Emancipation served to divide the liberals from the revolutionaries. The Autocracy seized its opportunity. Intransigent leaders, such as Chernishevski, were removed from the scene. The Sunday School movement for eliminating illiteracy was smashed by Imperial Decree (McLeish, 1974). The opposition forces were dispersed. Peasant revolts were put down to the sound of grapeshot and the screams of the peasants under the Cossack knout. The revolutionaries were driven underground. Many continued to work clandestinely for the overthrow of the regime. Perforce they were driven to terrorism. Any other activity was only marginally open to them. This was increasingly so after the failure of the Movement “To the People” of 1864-1873. The primary objective of civilizing the peasant was replaced by an “act of faith” in anarchist terrorism — that is, an appeal to the destructive “instincts” of the masses by presenting a model which would call forth their spontaneous response (Carr, 1937).

### *The Student Movement*

The earliest student groups originated in the 1830s as discussion circles. Between the end of the Crimean War (1856) and the *Manifesto of Liberation* (1861) many secret, illegal, sometimes revolutionary groups, sprang up at the Universities of Moscow, Kiev, and Saint Petersburg. These circles made contact through certain key figures (Pavlov, Bekman, Muravski, and others). On the basis of forbidden materials produced by Herzen in London, they worked out a common core of ideas, aspirations and techniques. The groups sprang up for no apparent reason; they flourished for a time, then



they disappeared without trace. Individual students provided continuity. Utopian programmes were elaborated and various alternatives were debated about how to achieve the ideal commonwealth.

One of the earliest was a group of Moscow students who called themselves the *Vertepniki*, the “cavemen” or “troglodytes” (Venturi, 1952). They met regularly with other young men — army officers and office workers — to discuss the burning questions of socialism, republicanism, and the social life of the Russian commune. Like most of these circles, they believed, with Herzen and Ogaryov representative of the repentant nobles, that the intellectual classes owed some kind of social debt to the peasantry. The circle discussed and prepared for an approach to the people as early as 1855. The “cavemen” were untypical only insofar as they belonged to the lower social classes. Several had family links with the religious schismatics (Raskolniki). Under the Tsar, dissenters were persecuted, placed under police surveillance and often outlawed. As a result, they provided fertile soil for revolutionary ideas (Brower, 1975).

According to a police report of 1858, the Moscow student group consisted entirely of ardent republicans, with a similar group of conspirators in the University of Kharkov. This group had passed from the state of abstract debate to the discussion of methods of abolishing the Autocracy. Already, they had a plan to organise and indoctrinate the peasant masses with revolutionary ideas. (Venturi, 1952).

The Kharkov circle formed immediately after the Crimean War. The founders were two students, Bekman and Muravski from Kiev, members of the lower nobility (Pichkurenko, 1954). The defeat of Russia in the Crimean War had been the occasion which led to the conspiracy. These students predicted that a complete overturn of Russian society was the only way to redeem the past and ensure a brighter future. With two other students, Yefimenko and Zavadski, members of the *raznochintsy* (“commoners” or “those outside the official ranks”), Bekman and Muravski constituted the nucleus of the group. By the end of 1856 they had recruited eight members.

At this point the group coalesced with the Paskvilny Komitet (“committee of slanderers”). There were five students in this group, all from the upper nobility. They were animated by the ideals, later caricatured by Turgenev in the character of the “nihilist” Bazarov, in the novel *Fathers and Children*. They dismissed all conventional beliefs, customs and ideas as obsolete. They claimed to believe in nothing, hence the label of “nihilism”. Like Bazarov, they combined this hostility to tradition with a single-minded devotion to “exact science.” This circle had one simple and clear aim—total revolution, which must begin with the abolition of serfdom:

We believed that Russia would arise that very year. All we needed to do was to publish a few thousand copies of some sort of manifesto and send our agents everywhere with it. If these plans were carried out we expected that one fine morning people would waken up to find a bloody revolution had broken out all over Russia. (Koz'min, 1930)

These fantasies were based on the palpable discontent of all classes. It was widely believed, even in official circles, that a peasant revolution was just around the corner. Many (following Bakunin) believed that this would occur “spontaneously”, the peasantry taking up their axes, reacting to

generations of injustice, frustration and oppression. On the other hand, the Kharkov circle believed that a necessary condition for revolution was the dissemination of revolutionary ideas by education and the printed word. They believed that the peasants represented the only social class with the necessary potential of revolutionary energy; in asking how to harness this energy, the Kharkov circle immediately confronted the problem of mass illiteracy and the means of overcoming it (McLeish, 1974).

The conditions of this period are those revealed in 1897, when the first national census discovered an illiteracy rate of 78.9%, reaching more than 90% in the countryside. In the total population of 125,680,682 only 25,569,585 could read. Most of the twenty-five million literates (which could signify merely the ability to write one's name) were relatively young; they were males; and they lived in the towns in the Polish provinces or in Great Russia. Of nearly 23 million children of school age only a quarter were literate (Ivanovich, 1906, 38-55). Even this degree of literacy represented decades of devoted work by liberal educators who had to combat strong government and Church interference. The treatment meted out to Count Leo Tolstoy and his school at Yasnaya Polyana is typical of the obscurantist attitudes to education which prevailed in official circles.

The Kharkov and Kiev groups coalesced on the basis of a common programme. Their ideas were derived from Herzen and his *Free Russian Press* in London. Under the influence of Herzen and the progressive history professor Platon Pavlov who visited Herzen abroad, they became republicans. The group made plans for the subversion of the Kiev garrison and for propaganda in the villages. However, their only successful activity was the distribution of a manifesto about the intolerable condition of Russian society revealed by the war in the Crimea.

The growing realisation of the overwhelming difficulties of organising a revolution caused the disruption of the group. According to Zavadski, the circle degenerated into gossiping factions. It was saved from total disintegration only by Bekman and Muravski who established a link with the Kiev circle in 1858. This circle was organised around Professor Pavlov of Kiev. On a visit to Herzen in London, he had picked up the idea of the Sunday School and of involving the peasants in the social movement by combining literacy education with revolutionary ideas (McLeish, 1974 and refs.).

The five students (later groups used five-member units as the basic "cell" of the conspiratorial network) Bekman, Muravski, Yefimenko, Portugalov, and Zavadski became itinerant evangelists of revolution at this point. They set up circles of helpers and sympathisers. These organized schools, lectures, libraries, and discussion circles. Other students and intellectuals not directly revolutionaries followed suit in setting up Sunday Schools. These were often used as centres of revolutionary propaganda (Abramov, 1900). More than 2,000 teachers were thus recruited. Between 1859 and 1862 more than 300 schools were created through local initiative. Liberal democrats and revolutionaries united in the common task (Pichkurenko, 1954). The government, seizing the excuse of the fires in St. Petersburg in May 1862 (claimed to be a signal for revolution), closed down all schools for illiterates and all reading rooms in the Empire and the Army (Kropotkin, 1899).



Subsequently, participants in the literacy movement were sentenced to exile, birching, loss of civil rights, and in one case "civil execution" followed by forty years imprisonment in a fortress (Pichkurenko, 1954, pp. 113-114).

The Sunday Schools were revived in the early 1870s and operated clandestinely in towns and villages. In 1873 there was the great educational movement "To the People" (*V narod*). Thousands of students, intellectuals, and revolutionaries left their homes, their studies, their occupations to go into the Russian countryside to teach the peasants about modern science, constitutional government, hygiene, and revolution. It was a mass movement of desperate protest, undertaken in a condition of euphoric enthusiasm. Like the Children's Crusade, to which it bore many analogues, the movement failed completely. Apathy and ignorance caused the village population to respond by driving the educators from the villages. Sometimes they informed the police, sometimes they simply boycotted these alien intruders. The only result of the movement was that many of the would-be educators were exiled or imprisoned.

The apathy of the peasants as a class was about equal to that of the liberals. Their political and social demands were limited in character. They were largely met in the economic consequences of Emancipation. The unrelenting pressure of the Tsarist police after 1873, the effects of the fanatical adventurism of Nechayev on the revolutionary movement, the debacle of the *V narod* policies, had their effect. Students decided that academic discussion and propaganda were luxuries they could ill afford if they involved payment by one's life, or a lifetime in a Tsarist dungeon (Lemke, 1908, 1923). The lack of response to the demands for constitutional reform, the success of police *agents-provateurs* in permeating and destroying the revolutionary groups, called forth a spirit of frustrated retaliation (Deacon, 1972, Chaps. 5-7).

These factors created a social vacuum where the violence of the state machine could be countered only by the violence of the terrorist. The revolutionaries, as they explained in proclamations after the assassination of Alexander II (Footman, 1944; Venturi, 1952) were driven to terrorism, in spite of their wishes. The consequences of the assassination of 1 March 1881 (or the lack of consequences) demonstrated the nature of the blind alley into which they had been driven. This blind alley was the terminus of the path they had taken under coercion of the political apathy of "the general will."

### *From Student to Professional Revolutionary*

In the revolutionary movement we can see a clear development. Initially the impetus is provided by adolescent romanticism and adventurism. The reaction from established authority forces a more realistic and scientific outlook on the movement. Control passes to the dedicated professional. Usually this individual is a victim of judicial persecution who develops a sublime faith in the outcome of the struggle and of his ability to bear with fortitude the extreme punishments the enemy has in store, or of his chances of avoiding them. A process of selection is at work: It is a matter of the "Survival of the fittest." Only the disciplined, organized, and dedicated revolutionary can hope to survive for more than eighteen months in the struggle against the secret police. This was the average "life-expectancy" in



the revolutionary movement. Chernishevski's fictional character Rakmetov, the hero of "What is To Be Done?" became the revolutionary ideal, a model for such as Vladimir Ilyitch Ulyanov (Lenin) three decades later. This hero is described by Nechaev:

The revolutionary is a dedicated man. He has no personal interests, no personal business, no sentiments, no attachments, no property, not even a name. Everything in him is absorbed by one single exclusive interest, one single thought, one single passion: the revolution. (Carr, 1937, Chap. 28)

The revolutionary was disgusted with Russian life and Russian conditions. His favourite characterisation of Tsarist society was as a graveyard. In his famous letter to Michelet, published in 1851, Herzen refers to Russia as "the deception, the pestilence".

This Russia extends from the Emperor, passing from gendarme to gendarme, from official to official, down to the lowest policeman in the remotest corner of the Empire . . . This living pyramid of crimes, abuses and bribery . . . is held together by a community of interest in plunder and gain, and supported on six hundred thousand animated machines with bayonets. (Herzen, 1927, Vol. 6, p. 236)

This description could be paralleled many times over from the works of the "Westernisers" Byelinski, Chernishevski, Mikhailov, Lavrov.

Rejection and hatred of the Tsar and the ruling class preceded any rational analysis of actual social conditions. The leading motif in political discussion is the question: "What force is there in Russia which can be counterposed to this utterly disgusting regime?" The intellectual, often a scion of the minor nobility, actually wanted to *become* a peasant or industrial worker, to devote his life to elevating the lower classes by educational or political means. Ogaryov is the prototype of this "repentant noble" (Carr, 1949). The denial of class origin was the main basis of the Student Movement of 1873—"To the People!" The comment of Rostopchin on the Decembrists, that it was the first time a revolution had been planned by gentlemen who wanted to become cobblers, was much more apposite in 1873 than it was in 1825.

### *The Value System and Activities of the Student Revolutionaries*

The theme of regeneration is fundamental: Revolutionaries agree that "Mother Russia" must be saved by the action of the millions of peasants. Regeneration can be effected only by destruction of the hated "White Father." It is not difficult to find an Oedipus theme in the writings and speech of the revolutionaries. There is a common desire, a religious wish, for internal peace, to be achieved by the return to some primeval Arcadia. The Paradise, lost by the design of wicked men, can be regained only at the expense of much blood and suffering. Some think of this Paradise as the Russian village, the commune, cleansed of the filth and disorder of centuries of malevolent intervention by alien intruders ("the Byzantine-German government"). Everything will be in common: All the needs will be fulfilled. Others see a Utopia emerging on the basis of industrial life, founded on science. Shchapov is a representative of this trend:

At all times and places ignorance of the laws of nature has produced only slaves; slaves of nature herself and of human force, of the force of muscle, brain and intelligence, of the force of deceit and prestige, wealth, power, despotism. (Shchapov, 1906-1908, Vol. 2, p. 170)

The seven stages characteristic of social change described by Parsons and Smelser can be easily recognised in the history of the Russian revolutionary movement (Smelser, 1959; Parsons & Smelser, 1957). Throughout the forties, fifties, and sixties, revolutionaries sought to add to the store of viable techniques and ideas. There were many changes in emphasis in thinking about the mechanics and anticipated progress of the revolution. These modifications and additions to revolutionary thought represent two stages in the learning process, proliferation of new ideas and attempts to specify the correct solution (Smelser, 1959, pp. 252-261). A few illustrations of these new ideas include: Serf emancipation; recruitment of dissenters to the revolutionary party; revolutionary terror (assassination, banditry); the effective organization of revolutionary groups; propaganda; a continuing process of education of the people in scientific knowledge; day-to-day agitation; literacy education and the saturating of it with revolutionary ideas; struggle for a constitution by legal means; each of these was proposed as a political panacea (Venturi, 1952).

The measures adopted by government represent marked symptoms of disturbance. In turn, these measures gave rise to disturbance in the opposition groups. These symptoms took the form of anxiety, violence, Utopian fantasies, neurosis, even insanity.

The numerous attempts to apply the method of choice resulted invariably in failure. Some of the most notable failures were: the assassination of Alexander II; the movement of liberalise the universities by direct action; the liberation from serfdom; the Russian Sunday Schools; the organization of trade unions and political parties.

The methods by the Autocracy to control innovation were highly successful. These were: exile to a distant province; civil execution; imprisonment for long periods, with hard labour and solitary confinement; a formal declaration by the Tsar that the innovator was insane; the use of *agents-provocateurs* by the Third Section; censorship of all published materials; the setting of one section of the radical intelligentsia against the other; general techniques of social disapprobation.

The effect of these measures was to slow down drastically the process of innovation. The sequence was: dissatisfaction gave rise to disturbance; disturbance in the system produced attempts at handling. In turn, the handling of symptoms showed the need for new ideas. Numerous attempts to specify then ensued. These different specifications were then tried. It is a matter of history that these trials invariably failed. This was the normal sequence of causal relationships (McLeish, 1968, pp. 65-66). The net outcome was that attempts to introduce a new order invariably broke down. Because of this abortive pattern, the spirit of innovation languished in Russia from 1790 to 1917. There were, of course, numerous social changes during this period, many extremely radical in nature. As examples, we may instance the emancipation of the serfs; the establishment of large concentrations of industry supported by foreign capital; Imperialist expansion, including the conquest of the Ukraine, Poland, the Caucasus, Armenia, Lithuania and Siberia — the mobilization of the population for the defeat of Napoleon; the debacle of the Crimean, then the Japanese war; the concentration of enormous debts to foreign money-lenders; the establishment of a State



Duma. The Russia of Nicholas II was not the same as that of Peter. But change was localized in particular sections of the economy and of the social system. The structure itself and the mass of the population were hardly affected.

The history of social change in Russia can be conceptualized therefore as a cyclical process with the process being pushed back repeatedly, the lost ground being regained by each generation of revolutionaries who rapidly recapitulated earlier stages. More concretely, there were a number of abortive attempts at change based on recurrent dissatisfaction with role-performance and with the use of material resources. This dissatisfaction arises afresh in each generation, due to the frustration of the needs of individuals and of defined groups. Building on the experience of their predecessors, each generation attempts to abbreviate the process of change so as to produce a successful outcome. They seek to begin from the point reached in previous sequences, to assimilate the heritage of revolutionary ideas as quickly as possible. The paradigm of trial and error learning can be used to explain social innovation in this case. Social change is a continuous process of adaptation to the exigencies of the environment.

The last two stages of innovation, namely, trial and success, followed by routinisation, appear only during the Soviet period.

#### *The Social Function of the Revolutionary Movement*

Speaking broadly, the Russian revolutionary movement served as the channel of communication between the most radical thinkers of Western Europe and the traditional culture of Tsarist Russia. It gave clear, if doctrinaire, expression to the needs and aspirations of certain classes or groups, in particular, the peasants, the sects, the Jews, the industrial workers. These groups had no possibility otherwise of communicating their sufferings, their hopes, their ideals except through passive resistance, or by unpredictable violence.

The intellectuals, active in the revolutionary movement, provided a voice for all the oppressed classes who had no influence on the course of events. They gave general expression, often in a language incomprehensible to their "constituents," to the need for change. Objectively speaking, they operated as a mediating influence to civilize the protest of the oppressed. In their absence, the reaction to oppression might otherwise have been Pugachovshchina—the murder of proprietors, the maiming of cattle, the burning of manors and manorial records, the flight from the village of whole populations, banditry. The revolutionary movement, appearing superficially to be a conflict phenomenon, performed a role useful to the established powers as well as to society as a whole.

It may seem paradoxical, but it is nevertheless true, that the Tsarist autocracy co-operated with the movement in a number of ways. This fails to be recognized, because in form this co-operation appeared as conflict. But by censoring writings, instead of totally suppressing them, official society was providing the revolutionaries with a licence to publish. The Tsarist police helped in a great number of ways to organize and "supervise" the work of the revolutionaries. We may instance Moscow police-chief Zubatov's labours in helping to organize the trade union movement. The secret police who



actively worked in the movement, through their reports, channelled ideas through to the Emperor himself. This meant that administrative decisions would normally take some account of the views of the clandestine as well as of the open publicist groups. The Emperor kept himself informed of the underground ideas by regularly reading the revolutionary newspapers produced by Herzen in London—he did not enquire too closely by what channels they arrived so promptly on his library table. At various times, particular sections of the revolutionary movement were encouraged at the expense of others. For example, Marxist groups were initially free to publish even the most radical views about society since their economic analyses and forecasts were in the form of a polemic against the Narodnik economists, then considered to be the most dangerous enemy of the established order. Later the “legal Marxists” were given similar support against the revolutionary wing. Certain Tsars, especially at the outset of their reigns, gave rise to expectations of reform. In this connection, a case could easily be made out for Peter the Great as the first pioneer of the revolutionary movement. Even the two Alexanders originally had philanthropic impulses.

#### References

- Abramov, Ya. D. *Nashi voskresniye shkoly; ikh proshloye i nastoyashchiye*. Spb., 1900.
- Brower, D. R. *Training the Nihilists*. Ithaca, N.Y.: Cornell University Press, 1975.
- Carr, E. H. *Michael Bakunin*. London: Macmillan, 1937.
- Carr, E. H. *The romantic exiles*. Penguin Books, 1949.
- Chekhov, N. V. *Typy russkoi shkoly*. Moscow, 1923.
- Deacon, R. *A history of the Russian Secret Service*. N.Y.: Taplinger Publishing Co., 1972.
- Footman, D. *Red Prelude: A life of A.I. Zhelyabov*. London, 1944.
- Herzen, A. The Russian people and socialism: A letter to J. Michelet in *My past and thoughts*, Vol. 6. London: Chatto & Windus, 1927.
- Ivanovich, V. (Chernolusski). Iz itogov pervoi russkoi vseoyobshchei perepisi. *Vestnik Vospitaniya*, 1906, 7, 38-55.
- Johnson, W. H. E. *Russia's educational heritage*. Pittsburgh: Carnegie Institute of Technology, 1950.
- Koz'min, B. P. *Kharkovskiye zagovorschchiki 1856-1858 godov*. Karkov, 1930.
- Koz'min, B. P. *Zaichnevski i 'Molodaya Rossiya'*. Moscow, 1932.
- Kropotkin, P. *Memoirs of a revolutionist*. London: Smith Elder & Co., 1899, 2 vols.
- Lampert, E. *Sons against fathers: Studies in Russian radicalism and revolution*. London: Oxford University Press, 1965.
- Leary, D. B. *Education and autocracy in Russia*. New York: 1919.
- Leites, N. *A study of Bolshevism*. Glencoe, Ill.: Free Press, 1953.
- Lemke, M. *Ocherki osvoboditel'nogo dvizheniya shestidesyatikh godov*. Moscow, 1908.
- Lemke, M. *Politicheshiye protsessy v Rossii 1860-kh godov (po archivnym dokumentam)*, Moscow-Leningrad, 1923.
- Lenin, V. I. *The development of capitalism in Russia*. Moscow, 1957.
- Lyashchenko, P. N. *Istoriya narodnogo khozyaistva SSSR*, 4th ed. Moscow, 1956, 2 vols.
- McLeish, J. *The theory of social change*. London: Routledge & Kegan Paul, 1968.
- McLeish, J. The Soviet conquest of illiteracy. *Alberta Journal of Educational Research*, 1972, 18, 307-326.

- McLeish, J. The Russian Sunday Schools — An experiment in voluntarism. *Alberta Journal of Educational Research*, 1974, 20, 75-95.
- Masaryk, T. G. *The Spirit of Russia*. London: 1919, 2 vols.
- Meijer, J. M. *Knowledge and revolution: The Russian colony in Zurich*. Assen: Van Gorcum, 1955.
- Parsons, T. & Smelser, J. *Economy and society*. London: Routledge & Kegan Paul, 1957.
- Pichkurenko, Ya. D. K voprosu o roli voskresnikh shkol v burzhuaz-nodemokratisch'eskom osvoboditel' nom dvizhenii v Rossii v kontzye 50-kh k nachalye 60-kh godov XIX v. *Sovetskaya Pedagogika*, 1954, 5, 96-114.
- Ralston, W. R. S. *The songs of the Russian people as illustrative of Slavonic mythology and Russian social life*. London: Ellis, 1872.
- Shchapov, A. P. *Sochineniya*. Spb., 1906-1908, 3 vols.
- Smelser, N. J. *Social Change in the Industrial Revolution*. London: Routledge & Kegan Paul, 1959.
- Venturi, F. *Il populismo russo*. Firenze, 1952, 2 vols. (English trans. by Francis Haskell, 1960).
- Wallace, D. M. *Russia*. London: Cassell, 1912.

E. A. BLOWERS

*The University of Alberta*

## Prediction of Metropolitan Readiness Test Scores

*Readiness tests are used widely in kindergarten and grade one as useful predictors of future school performance. In this study, the efficiency was examined of several visual and auditory predictors of the Metropolitan Readiness Tests. Subjects were 106 grade one children considered by their teachers to show learning difficulties. The stepwise multiple regression analysis of the data indicated that visual tasks were more efficient as predictors. It was suggested that the Metropolitan may be overly dependent on visual abilities and that some development of abilities considered as efficient predictors of readiness scores might be done in kindergarten programs. (Dr. Blowers is Assistant Professor in the Department of Educational Psychology, The University of Alberta.)*

Readiness tests are used widely in kindergarten and grade one classes to determine the general aptitude of young children for academic learning. One of the better known readiness tests is the Metropolitan (Hildreth & Griffiths, 1955) which is administered to all grade one pupils throughout cities such as Washington (Blackmore, 1973) and Edmonton (Edmonton Public School Board, 1975). *The Seventh Mental Measurements Yearbook* (Buros, 1972) describes the Metropolitan as an appropriate predictor of early formal learning. In summary, the Metropolitan Readiness Tests are used widely and are regarded as effective in predicting early school learning. This study examines the question of whether other types of tests commonly used with preschool and grade one children may be useful in predicting scores on the Metropolitan Readiness Tests.

### *Method*

Fifty female and 56 male subjects ranging in age from 67 to 78 months and attending normal grade one classes in the public school system of a



large Canadian urban centre were selected by their teachers as possibly experiencing learning difficulties. While it is not possible to judge how representative the group selected was of all grade one pupils with learning problems, all teachers who were asked to take part in the study identified one or more of their pupils as appropriate for the study. The subjects in the

TABLE 1  
INSTRUMENTS USED IN THE STUDY

Category	Test	Subtest	Content	Method of Presentation	Method of Response
Readiness	Metropolitan	Word Meaning	Relating words to pictures	Auditory	Visual-motor, marking appropriate line drawing
		Listening	Relating sentence/paragraph to pictures	Auditory	Visual-motor, as above
		Matching	Locating shape/design/word identical to stimulus	Visual	Visual-motor, as above
		Alphabet	Selection of letter named	Auditory	Visual motor, marking appropriate letter
		Numbers	Range of pre-arithmetic and early arithmetic skills including comparison, counting basic number operations	Auditory	Visual-motor, as in Word Meaning
		Copying	Copying shapes/designs/letters	Visual	Visual-Motor
Auditory	Goldman-Fristoe-Woodcock Discrimination	Silent	Discrimination of acoustically similar speech sounds	Auditory-from tape recording	Visual-motor, selection of appropriate picture
	Detroit Tests of Learning Aptitude	Related Syllables	Meaningful auditory memory for sentences	Auditory-from tape recording	Auditory-Vocal subject repeats sentence
	Peabody Picture Vocabulary Test		Receptive vocabulary-relating words to pictures	Auditory	Visual-selection of appropriate drawings
	Illinois Test of Psycholinguistic Abilities	Sequential	Nonmeaningful auditory memory for digits	Auditory	Auditory-vocal
		Auditory Closure	Recognition of partially heard words	Auditory	Auditory-vocal
		Sound Blending	Combining speech sounds to form words	Auditory	Auditory-vocal
Visual	Illinois Test of Psycholinguistic Abilities	Visual Memory	Recall and reproduction in sequence of nonmeaningful visual stimuli	Visual	Visual-motor
		Visual Closure	Recognition of partially visible stimuli	Visual	Visual-motor
	Frostig Developmental Test of Visual Perception	Eye-motor coordination	Construction of relatively straight lines between two points	Auditory and visual	Visual-motor
		Figure-ground Discrimination	Outlining overlapping and/or obscured figures	Primarily visual	Visual-motor
		Form constancy	Recall and discrimination of common shapes	Auditory and visual	Visual-motor
		Position in space	Shape matching by fine directional cues	Visual	Visual-motor
		Spatial Relations	Copying marble board designs on opposite side of page to which presented	Visual	Visual-motor

study, therefore, represented a group of pupils considered by their teachers to be displaying some difficulty in mastering grade one academic work.

All subjects received the Wechsler Intelligence Test for Children (Wechsler, 1955) and ranged in full scale scores from 69 to 131, with a mean score of 102. The socioeconomic backgrounds of the families ranged from 29 to 76 on the Blishen Occupational Class Scale (1958).

Predictor variables used in the study are in relatively common use with kindergarten and grade one children and included parts of the Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy & Kirk, 1968), the Frostig Developmental Test of Visual Reception (Frostig, Lefever & Whittlesey, 1963), the Peabody Picture Vocabulary Test (Dunn, 1959), the Goldman-Fristoe-Woodcock Test of Auditory Discrimination (1970) and the Memory for Related Syllables subtest of the Detroit Tests of Learning Aptitude (Baker & Leland, 1959). The predictor variables were classified as primarily auditory or primarily visual. Auditory predictors included measures of abilities involved in discrimination, receptive vocabulary, meaningful and nonmeaningful sequential memory, closure and sound blending. Visual predictors included measures of abilities involved in sequential memory, closure, eye-motor coordination, outlining obscure and overlapping figures, shape discrimination, shape matching by directional cues and reproducing pictured marble board type designs.

Predictor variables were classified as primarily auditory or primarily visual in relation to the content of the test or subtest. At the same time, the presentation of the material and the response required of the subject might be primarily auditory, primarily visual or some combination of these. To take one example, the Word Meaning subtest of the Metropolitan is presented orally and the subject is required to select a picture that illustrates the word by marking the picture. This indicates that the Word Meaning subtest is a relatively complex task involving auditory vocabulary and memory as well as visual discrimination and some level of visual motor ability. In general, many of the tests used tended to draw upon a variety of abilities. The instruments used in this study are described in Table 1 by title, content, and methods of presentation and response.

The statistical analysis was designed to identify those variables best predicting subtest and total Metropolitan scores. Stepwise multiple regression analysis was used to determine which combinations of predictors were most efficient in predicting readiness test scores (Ferguson, 1971). In making the stepwise multiple regression analysis, all predictor variables were included to determine which were most efficient in predicting the criteria variables. Predictors at a probability level of .05 or less were considered relatively efficient.

### *Results and Discussion*

A summary of significant predictors at the .05 level of significance, with the next most efficient predictor above the .05 level, is given in Table 2.

The level of prediction of the subjects of the Metropolitan Readiness Test varied from approximately 54% in Numbers to only 27% in Listening. Seventy percent of the variability in the total score was accounted for by variability in four of the predictors. At the probability level of .05 or less,

TABLE 2  
STEPWISE MULTIPLE REGRESSION ANALYSIS PREDICTION OF  
METROPOLITAN READINESS TEST SCORES

Criterion Variable	Predictor Variable	F	Probability	Cumulative % of Variance
Word Meaning	Receptive Vocabulary	38.43	0.00	26.98
	Meaning Auditory Memory	9.68	0.00	33.26
	Eye-motor Coordination	3.40	0.07	35.45
Listening	Marble Board Type Designs	23.55	0.00	18.46
	Receptive Vocabulary	11.47	0.00	26.63
	Auditory Discrimination	3.42	0.07	29.01
Matching	Marble Board Type Designs	27.17	0.00	20.71
	Shape Matching	4.92	0.03	24.33
	Shape Discrimination	2.44	0.12	28.87
Alphabet	Marble Board Type Designs	36.64	0.00	20.05
	Receptive Vocabulary	7.99	0.01	31.38
	Nonmeaningful Auditory Memory	5.12	0.03	34.66
	Shape Discrimination	3.26	0.07	36.70
Numbers	Marble Board Type Designs	55.81	0.00	34.92
	Meaningful Auditory Memory	27.88	0.00	48.79
	Receptive Vocabulary	6.27	0.01	51.75
	Shape Matching	4.70	0.03	53.89
	Auditory Closure	2.90	0.09	57.51
Copying	Marble Board Type Designs	38.44	0.00	26.99
	Shape Discrimination	7.67	0.01	32.05
	Eye-motor Coordination	5.35	0.02	35.43
	Receptive Vocabulary	3.73	0.06	37.73
Total Scores	Marble Board Type Designs	44.02	0.00	47.48
	Receptive Vocabulary	41.80	0.00	62.64
	Shape Discrimination	17.18	0.00	68.03
	Meaningful Auditory Memory	8.94	0.00	70.63
	Auditory Closure	3.32	0.07	71.57

auditory instruments were most effective in predicting only Word Meaning, while visual instruments were most effective in predicting Matching and Copying. The single best predictor of most of the Metropolitan subtest scores and of the total scores was the Marble Board Type Designs, subtest five of the Frostig Developmental Test of Visual Perception, which appears to depend upon visual-motor abilities both in presentation and response (Table 1). The subtests of Copying and Matching appear to depend primarily on visual-motor abilities but the best single predictor of the Listening subtest also was Marble Board Designs. All subtests on the Metropolitan require interpretation of symbolic visual material, which may explain in part why visual and visual-motor abilities appear to influence the scores on this test. It is possible that the Metropolitan Readiness Test is less useful than it might be in predicting academic learning because of this limited influence of auditory abilities on scores. To take an example, a child with high visual but low auditory abilities might appear to have a spuriously high chance of success in early school activities, as his readiness test scores might be only minimally related to his problems.



References

- Baker, H., & Leland, B. *Detroit tests of learning aptitude*. Indiana: Bobbs-Merrill, 1959.
- Blackmon, A. Early childhood education in the public schools of the nation's capital. In B. Fallon (Ed.), *40 innovative programs in early childhood education*. Belmont, California: Fearon, 1973.
- Blishen, B. The construction and use of an occupational class scale. *Canadian Journal of Economics and Political Science*, 1958, 24(4), November, 523.
- Buros, O. *The seventh mental measurements yearbook*. Highland Park, N.J.: Gryphon Press, 1972.
- Dunn, L. *The Peabody picture vocabulary test*. Minneapolis: American Guidance Service, 1959.
- Edmonton Public School Board. *Current system and optional measurement program*. Edmonton: Unpublished, 1975.
- Ferguson, G. *Statistical analysis in psychology and education* (3rd ed.). New York: McGraw-Hill, 1971.
- Frostig, M., Lefever, D., & Whittlesey, J. *The Marianne Frostig developmental test of visual perception*. Palo Alto: Consulting Psychologists Press, 1963.
- Goldman, R., Fristoe, M., & Woodcock, R. *Goldman-Fristoe-Woodcock test of auditory discrimination*. Circle Pines, Minnesota: American Guidance Service, 1970.
- Hildreth, G., & Griffiths, N. *Metropolitan readiness tests*. New York: Harcourt, Brace & World, 1955.
- Kirk, S., McCarthy, J., & Kirk, W. *The Illinois test of psycholinguistic abilities*. Urbana: University of Illinois Press, 1968.
- Wechsler, D. *Wechsler intelligence scale for children*. New York: Psychological Corporation, 1955.

## BOOK REVIEW

THE FORKED ROAD: CANADA 1939-57. By Donald B. Creighton. Toronto: McClelland and Stewart, 1976

In the Canadian Centenary Series of which he is also advisory editor, Donald Creighton, one of Canada's senior and most productive historians, has contributed in the *The Forked Road* a perceptive history of Canada covering the eighteen years from 1939 to 1957, years dominated nationally by Liberal governments successively under the leadership of the strange but politically astute Mackenzie King and the urbane, business-like but rather unimaginative Louis St. Laurent.

Having first recorded the entry of Canada into World War II, Creighton takes a cross-sectional look at us as we were in 1939: in politics, international relations and policies, technological and social development, religion, education, and the arts. Creighton's thesis, as suggested in the title of the book, is that Canada, having arrived at a point in her history where two roads to the future diverged, chose the one that "led directly into the new world of planning and management, of economic controls and social equalization." This new world was marked by a weakening of ties with Great Britain and the Commonwealth and by increased dependence on and cooperation with the United States. It was marked, too, by emphasis on the use of advanced technology in the development of both primary and secondary industries. Manufacturing began to flourish as never before; natural resources, including expanding sources of petro-chemical and electric energy, gave new direction and impetus to our physical and economic growth. All of this affected the demand for consumer goods, the general standard of living, and the increasing use of the means of communication and travel. To make these possible required bold, if sometimes questionable, decisions involving large amounts of capital secured, in increasingly large amounts, through foreign investment, international loans, and government spending.

In spite of Mackenzie King's efforts to prevent conscription and to limit Canada's contribution as much as possible to a training (e.g., Commonwealth Air Training Plan) and supply role, Canadians were proud of the achievement of their army, navy and air forces in various of the war zones. While she played a minor part in determining strategy and tactics related to the conduct of the war and in arriving at a peace settlement, Canada soon began to participate as an important middle power in affairs of the newly created United Nations, in the North Atlantic Treaty Organization, and in

the solution of such international problems as those related to the partition of Palestine, the Korean War, and the closure of the Suez Canal. On the international scene, Lester B. Pearson, who had been associated with the Department of External Affairs for many years, became increasingly prominent as he moved from civil servant to full parliamentary responsibilities.

Some goals of the government were not achieved during the long period of Liberal rule: patriation of the constitution, agreement on a distinctive Canadian flag, and the cementing of Canada into a truly unified nation. Other goals, however, met a better fate: Canadian citizenship became a fact, appeals from Canadian courts to the Judicial Committee of the Privy Council were abolished, for the first time a Canadian (Vincent Massey) became Governor-General, Newfoundland became the tenth Canadian province, family allowances were introduced, and old age pension provisions were broadened. The concept of social welfare was becoming well established.

Through both private and government encouragement, the arts were beginning to receive greater recognition and support. An important step taken toward the end of the period was the creation of the Canada Council.

During the final years of the period, too, the government supported C. D. Howe, whose contributions during and following the war had been substantial, on two major issues, the passage of the Defence Production Bill and the Trans-Canada Pipeline Agreement. These were contentious matters which, in part at least, were factors in the defeat of the Liberals by John Diefenbaker and his Progressive Conservative Party in 1957.

Readers of the *Alberta Journal of Educational Research* will be interested in statements made by Creighton about education. After a brief description of the extent and nature of education in Canada in 1939, the author states realistically that "The system was entirely appropriate for those who were going on to the universities and the professions; but it was less than satisfactory for those whose education would end with high school or collegiate institute and who would then have to find jobs in an increasing industrialized world." A rapid increase in the birth rate following the war, coupled with a broadened curriculum at the secondary school level, led to growing school enrolments during the fifties. The return of veterans whose education had been interrupted by war and for whose rehabilitation financial incentives were provided led to the large number of mature and highly motivated students who entered universities, colleges and technical institutes in the late forties. An interest in graduate studies led to the provision of further facilities and opportunities.

The organization, the breadth of understanding, the attention to detail, the attractive style and flow of expression make this a compelling book for those interested in developments in Canadian history during the critical years of World War II and the post-war period, 1934-57.

H. T. Coutts  
*Dean Emeritus of Education*  
*University of Alberta.*



# EDUCATION IN CANADA

The Canadian Society for the Study of Education invites you to explore issues in education through its publications.

## CANADIAN JOURNAL OF EDUCATION / REVUE CANADIENNE DE L'EDUCATION

Volume 1 (1976) ..... \$8.00 (4 issues, \$2.50 each)

Volume 2 (1977) ..... \$10.00 (4 issues, \$3.00 each)

### *Some articles from Volume 1*

Livingston/ Images of the Educational Future in Advanced Industrial Society.

Fisk/ Branding Time at Bar-Alma U.

Gillett/ Self-Disclosure and the Women's Movement.

### *First issue in Volume 2. Special Issue on Canadian Studies.*

Robert/ Les conditions préalables dans le développement des études du Canada.

Duckworth/ Assessing the Canada Studies Foundation, Phase 1: An Approach to a National Evaluation.

## 1976 YEARBOOK : BILINGUALISM AND CANADIAN EDUCATION / LE BILINGUISME DANS L'EDUCATION CANADIENNE

( Merrill Swain, editor) ..... \$3.00 each

### *Some articles in the 1976 Yearbook*

Churchill/ Recherches récentes sur le bilinguisme et l'éducation des francophones minoritaires au Canada.

Greenfield/ Bilingualism, Multiculturalism and the Crisis of Purpose in Canadian Culture.

*Discounts: 40% on 5 copies or more.*

Canadian Society for the Study of Education / Société  
canadienne pour l'étude de l'éducation

Faculty of Education  
University of Alberta  
Edmonton, Alberta T6G 2G5









## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centred and in capital letters, e.g., TABLE 4. The heading of the table is to be centred below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. the position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last name at the end of the manuscript under the heading, *References*.
6. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
7. In matters of style, the APA Publication Manual is considered definitive.





# ajer

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXIII    NUMBER 3    SEPTEMBER 1977

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA    ●    EDMONTON

## THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

.....

Published quarterly in March, June, September and December  
by the  
Faculty of Education  
of  
The University of Alberta

.....

### Faculty Publications Committee

*K. A. Neufeld*, Department of Elementary Education  
*J. Chambers*, Department of Educational Psychology  
*A. K. Clark*, Department of Industrial and Vocational Education  
*A. G. Konrad*, Department of Educational Administration  
*M. R. Lupul*, Department of Educational Foundations  
*R. G. Martin*, Department of Secondary Education  
*C. H. Preitz*, Department of Industrial and Vocational Education

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

Chairman: *K. A. Neufeld*

Editor: *Andrew K. Clark*

Secretary: *Mrs. A. Onishenko*

.....

AJER gratefully acknowledges support from Canada Council and AACES. AJER is indexed in the *Canadian Education Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, and *Sociology of Education Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available in bound and unbound volumes; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, The University of Alberta, Edmonton, Alberta, Canada T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



# The Alberta Journal of Educational Research

Volume XXIII, No. 3

September, 1977

## CONTENTS

Children's Ability to Learn Problem-Solving Strategies from Television .....	171
<i>T. V. Hodapp</i>	
Young Children's Behavior in Solving Division Problems .....	178
<i>R. Bourgeois and D. Nelson</i>	
Beginning Readers Discriminate Words by Their Simplest Cue .....	186
<i>J. Hundert and B. Bucher</i>	
A Naturalistic Study of Reciprocity in the Helping Behavior of Young Children .....	195
<i>R. F. Marcus and B. Jenny</i>	
Differential Attitudes Toward Mainstreaming: An Investigation .....	207
<i>R. J. Williams and B. Algozzine</i>	
A Critical Examination of Dewey's Method of Inquiry .....	213
<i>H. W. Hodysh</i>	
Equality of Educational Opportunity and Inequality of Per Pupil Expenditures .	218
<i>L. M. Bezeau</i>	
Schools I would Like to See: An Opinion Survey Instrument with Interesting Possibilities .....	226
<i>J. C. Powell, D. J. Cottrell and M. Lever</i>	
BOOK REVIEW	
<i>Teaching Elementary School Mathematics Through Motor Learning,</i>	
R. B. Ashlock and J. H. Humphrey .....	242
Reviewed by D. Sawada	

FACULTY OF EDUCATION  
*The University of Alberta*





*Announcement  
of the*

**Faculty of Education  
Lecture Series Award**

The Faculty of Education of the University of Alberta inaugurated in 1977 a Lecture Series Award as one means of encouraging, supporting and publicizing educational research. The recipient of this award will be granted a six-month research leave commencing July 1, 1978. During this time the award holder will be expected to engage in the pursuit of his/her research activities and, as well, prepare a three part lecture series which will be presented in January 1979. At least one of these lectures will deal explicitly with the significance of the research for teaching and/or teacher preparation.

This award is intended to identify and reward those members of the Faculty of Education who have been engaged in educational research to the point where they have:

- (a) gained a reputation among leading scholars as a significant contributor to the advancement of knowledge in their field, through the development and pursuit of a sustained program of research,
- (b) disseminated their findings through the presentation of papers, and/or the publication of articles in refereed journals or books,
- (c) identified a worthwhile research undertaking which can be both significantly advanced and reported as a result of receipt of the award.

**ELIGIBILITY:**

Only full-time tenured members of the Faculty of Education of the University of Alberta are eligible for the award.

**DEADLINE FOR APPLICATION OR NOMINATION:**

The Advisory Committee on Selections will commence its work of compiling and reviewing dossiers October 1, 1977. Therefore, nominations of colleagues or of self should be directed to the Chairman of the Committee prior to September 30, 1977. No supporting documentation is required with the nomination.

The research leave will not necessarily be awarded annually.

Advisory Committee,

Juanita Chambers  
Les Gue  
Bernard Schwartz  
Robert Patterson (Chairman)

# new from Manchester

TEACHING LOCAL HISTORY W B Stephens Designed mainly for teachers and student teachers at junior, middle and secondary school level, this book discusses the case for the inclusion of local history in the curriculum. The practical difficulties are examined, and suggestions made for work both in and out of the classroom. There is much information on source material, and the use of printed and pictorial matter, as well as a substantial section on field-work and archaeology. Illustrated £5.95 net

A RAGGED SCHOOLING Growing up in the Classic Slum Robert Roberts 'Moving and enthralling ... filled with humour, hope and human warmth. The finest and truest picture of the other side of the Edwardian era.' - Daily Telegraph 'The autobiography of an exceptional man ... quite extraordinary.' - Observer 'Marvellous ... Roberts watched, observed, marvelled - and remembered.' - Spectator Illustrated £3.75 net

PHILOSOPHY OF EDUCATION A SELECT BIBLIOGRAPHY Compiled by John P Powell The object of the bibliography is to provide a guide to the rapidly growing and somewhat scattered literature on the philosophy of education. The third edition of this important reference work is now available; it contains many new items and has extended its classification scheme to include many previously neglected areas of enquiry. Each item is numbered for cross-reference and there is a full author and journal index. paper £2.25 net

RESEARCH IN EDUCATION Edited by A D Edwards and J Rushton The main emphasis of the journal is on empirical research in all the recognised fields of education, together with interdisciplinary studies; and includes papers on research in education at all levels. There is also a regular book review section. The journal is published twice yearly. 1978 subscription \$15, single copies \$8

THE TEACHER IN A CHANGING SOCIETY Edited by John D Turner and James Rushton £3.50 net

EDUCATION AND DEPRIVATION Edited by James Rushton and John D Turner £3.95 net

EDUCATION FOR THE PROFESSIONS Edited by John D Turner and James Rushton £4.50 net

MANCHESTER UNIVERSITY PRESS Oxford Road, Manchester M13 9PL



T. V. HODAPP

*ACCESS Alberta*

## Children's Ability to Learn Problem-Solving Strategies from Television

*A study was undertaken to determine whether young children can learn problem-solving strategies from educational television segments. Fifty-two 5.5 to 6.5 year old children were shown a segment illustrating a problem and its solution. Thirty-seven children of the same age were shown a control segment. Their eye movements were monitored as they viewed the segments to determine if they extracted relevant information from the segment. They were then asked to perform an identical or a transfer problem task. Results indicated that young children could imitate, but not learn and transfer problem-solving strategies from media to real life. Those who did not solve the problem task did not comprehend the TV segment in terms of a problem and its solution. (Dr. Hodapp is a Senior Research Officer at ACCESS Alberta, Edmonton.)*

Educational television stations and production corporations have begun developing programs which instruct young children in such skills as problem-solving strategies on the assumption that young children can abstract complex conceptual principles from media presentations. Studies have indicated that young children can solve simple problems by using strategies (Scott, 1973). However, opinions differ regarding the ability of young children to learn problem-solving strategies from media presentations and transfer them to real-world situations.

Piaget (1970) states that the transductive nature of young children's thought precludes skill learning from media. Olson and Bruner (1974) state that conceptual skills can only be acquired through repetition coupled with feedback from an expert.

Salomon (1972), on the other hand, states that cognitive processes can be replaced, or supplanted, by representations via media of more appropriate or higher order processes. Supplantation compensates for what the individual cannot learn on his or her own by providing an image of a transformation which is internalized and stored for later use. This is accomplished through

the overt execution of a process directly in front of the learner, causing by direct presentation what would otherwise have required direct manipulation by the individual.

Other research tends to support Salomon. Bandura, Ross and Ross (1963) found that young children will imitate televised models while Sullivan (1967) found that young children could learn conservation problems via film. Yussen (1974) found that in order for an action or a skill to be successfully imitated or transferred, the relevant aspects of the original situation must be observed.

Research on problem-solving (Gagne, 1966) has indicated that problem solution, i.e., overcoming a barrier to a goal, requires integration of previous experience with the requirements and limitations of the problem situation into a strategy of attack. Past experience may be the most crucial factor as it determines how the individual perceives the problem situation as a whole and all of the relevant aspects of the problem.

Whether a young child can observe the relevant aspects of a problem situation, much less a televised display embellished with devices designed to attract and maintain the attention of the viewer (Lesser, 1974), and apply what has been seen to a later situation is highly questionable. Research on the eye movements of young children (Vurpillot, 1976) indicates that young children tend to concentrate on the perceptually dominant features of a visual display. It is only after perceptual/cognitive structures have developed to the operative level that young children can observe in a coordinated fashion from a visual display information relevant to a specific question or problem (O'Bryan & Boersma, 1971).

Thus, the objective of this study was to determine whether young children will view the relevant aspects of a televised segment illustrating a problem-solving strategy in an entertaining fashion and then utilize this information as past experience for a subsequent problem-solving situation.

### *Method*

#### *Subjects*

The subjects for this experiment were 51 males and 38 females ranging in age from 5.5 to 6.5 years of age. All were residents of a metropolitan suburban community and attended kindergarten classes in one of two schools. Subjects were randomly selected from all kindergarten students in both schools. The SRA Short Test of Educational Ability was administered to all subjects in their classes prior to their participation in this study as a measure of IQ.

#### *Procedure*

The eye movements of each child were monitored as they viewed one of two problem-solving videotaped segments. The experimental segment illustrated how to obtain an object out of arm's reach by using a tool (a hooked object in this case). This strategy was deemed appropriate for subjects of the age employed by pilot testing. The 100-second segment was produced by the researcher and employed live actors and puppets in a game-like situation. The 110-second control segment was produced by the Children's Television Workshop for the program "Sesame Street" and illustrated the solution of a classification task.



After viewing the appropriate segment each subject was taken to a separate room and asked to solve one of two problems. The identical problem was literally identical to that demonstrated in the experimental segment in terms of setting and situation. Several objects, some also used in the segment, were located about the area as potential tools, but not in such a way as to make their use as an aid to problem solution obvious. The transfer task employed the same strategy as the identical task but with different objects in a different setting. Again, objects were placed about for use as possible tools. Subjects were scored for their times-to-solution by an assistant viewing behind a one-way mirror according to a predetermined criteria for problem solution. The method of solution was also noted if solution was attained within the 5-minute limit provided each subject.

A 2 x 2 (Segment x Task) design was employed in this study. Subjects were randomly assigned to one of the four treatment groups. The number of subjects in each group was: Experimental Segment-Identical Task = 29; Experimental Segment-Transfer Task = 18; Control Segment-Identical Task = 23; and Control Segment-Transfer Task = 19.

In this experiment a Polymetrics Products Eye Movement Recorder employing the principle of corneal reflection was employed to obtain videotaped recordings of the eye movement patterns. The stimulus material was presented to the child by means of a Sony 9" television monitor connected to a Sony ½" videotape recorder. Placement of the stimulus material insured maximum use of the viewing area of the screen.

A bright spot detector was attached to the video line. The first spot in the field of view (the corneal reflection) to exceed a brightness threshold was detected by the device and transmitted in terms of the length of the synchronizing pulse from the beginning of the visual field to the location of reflection to a PDP-9 mini-computer where it was transformed to a coordinate point.

The points of fixation thus determined were compared to the location of objects (targets) relevant or irrelevant to the problem-solving strategy in the visual field during the course of the problem-solving segments. Several measures were employed: number of times each target was viewed, percentage of total number of fixations on target, and the percentage of the total time on target.

### *Results*

Results of a 2 x 2 chi-square analysis indicated that frequency of solution of the identical task was significantly related to viewing the experimental segment ( $\chi^2 = 20.66$ ,  $df = 1$ ).

A 2 x 2 (Segment x Task) ANOVA with age and IQ as covariates was performed with time-to-solution as the dependent variable. The main effect for Segment was significant ( $F(1,28) = 8.56$ ,  $p < .01$ ) with the experimental segment group having the lower mean. The covariate age was also significant ( $F(1,28) = 6.94$ ,  $p < .01$ ). These results indicate that while the two tasks were relatively equal in difficulty, at least for those who could solve the problems, young children can only imitate problem-solving strategies viewed via television in identical situations and not transfer them to novel situations. The significant age covariate is accounted for by a significant



TABLE 1  
PERFORMANCE ON FILM IDENTICAL AND TRANSFER TASKS

	Solution Attained	Experimental	Control	
<u>Identical Task</u>	No	7	21	28
	Yes	22	2	24
		29	23	52
	$\chi^2 = 20.66, df = 1, p < .001$			
<u>Transfer Task</u>	No	11	13	24
	Yes	7	6	13
		18	19	37
	$\chi^2 = .01, df = 1, NS$			

TABLE 2  
ANALYSIS OF COVARIANCE: TIME-TO-SOLUTION BY FILM  
AND TASK WITH IQ AND AGE

Source	Sum of Squares	df	Mean Square	F
Age	28,993.430	1	28,993.430	6.940**
IQ	79.810	1	79.810	.019
Film	35,761.129	1	35,761.129	8,560**
Task	15.608	1	15.608	.004
Film Task	16.674	1	16.674	.004
Residual	116,975.375	28	4,177.691	
Total	256,683.812	33	7,778.297	

\*\*  $p < .01$

positive Pearson correlation between age and time-to-solution in the Identical-Experimental group. ( $r = .57, p < .01$ ). This was the only significant correlation to emerge between age or IQ and time-to-solution in any of the treatment groups.

Results of a 2 x 2 chi-square analysis indicated that, for those who solved the problem-solving task, viewing the experimental segment, as opposed to the control segment, was significantly related to the use of a hook as the solution tool ( $\chi^2 = 7.54, df = 1, p < .01$ ), as opposed to any other possible solution. This indicates further that children were able to imitate the solution strategy demonstrated in the experimental segment.

Results of another 2 x 2 chi-square analysis were not significant ( $\chi^2 = 1.45, df = 1$ ) which indicated no relation between segment viewed and type of behavior employed by nonsolvers in their attempts to solve the problem-

TABLE 3  
SOLUTION AND NONSOLUTION BEHAVIOR

	Experimental	Control	
<u>Solution Behavior</u>			
Hook	28	5	33
Other	1	3	4
	29	8	37
$\chi^2 = 7.54, df = 1, p < .01$			
<u>Nonsolution Behavior</u>			
Tools	2	1	3
Motor	16	33	49
	18	34	52
$\chi^2 = 1.45, df = 1, NS$			

TABLE 4  
SUMMARY OF REGRESSION ANALYSES

Variable	Pearson $r$	Beta Weight
Age	.55	.85**
Viewed hooked tool	.42	.50*
% Fixations--hooked tool	.59	.61**
% Time on Target--hooked tool	.53	.53**
% Fixations--hooked tool	.94	.89**
% Time on Target--hooked tool	.94	.99**
* $p < .05$		
** $p < .01$		

solving task. Ninety-four percent of the nonsolvers employed only motor behavior, as opposed to tools, in their attempts to solve the problem.

Comparisons using t-tests were conducted between problem task solvers and nonsolvers who had viewed the experimental segment. Separate analyses were performed for those who attempted the identical or transfer tasks. Dependent variables were the number of times each target was viewed during the segment, the percentage of the total number of fixations on target, and the percentage of the total time on target. No significant differences were found between problem-task solvers and nonsolvers in the analyses for either the identical or the transfer tasks. This indicates that the eye movements for solvers and nonsolvers were the same on the experimental segment.

Multiple regression analyses were conducted between time-to-solution, as

the dependent variable, and each of the three eye movement variables for those who viewed the experimental segment and solved the problem task. For those who solved the identical problem task significant predictors of times-to-solution were: age ( $r = .55$ ), frequency of viewing the hooked tool used to solve the problem during the segment ( $r = .42$ ), percentage of fixations on the hooked tool in an early scene ( $r = .59$ ) and percentage of time on the hooked tool in an early scene ( $r = .53$ ). For those who solved the transfer task, significant predictors of time-to-solution were: percentages of fixations on the hooked tool during an early scene ( $r = .94$ ), and high percentages of time on the hooked tool during an early scene ( $r = .94$ ).

### Discussion

Results of the above experiment indicate that children can imitate a problem-solving strategy from a televised segment illustrating that strategy. However, children were unable to abstract the strategy and apply it to a transfer task. It may be that children can learn strategies from television but cannot recall the strategy in the face of overwhelming perceptual differences between the televised demonstration situation and their actual problem situation.

Subjects who viewed the experimental segment and solved the problem task employed the same strategy as that modelled in the segment. Those who did not solve the problem task did not use tools in their attempts to solve the problem.

Analysis of eye movement measures indicated that nonsolvers did not differ in their viewing patterns from solvers. However, several of the eye movement measures on strategy related targets were predictive of time-to-solution.

This leads to the conclusion that those subjects who solved the identical task, and possibly the transfer task, comprehended the experimental segment in terms of a problem and a strategy for the solution to that problem. Those who did not solve the tasks had only a very superficial comprehension of the conceptual content of the segment.

The results reported here suggest an intermediate ground between Salomon and Olson, at least for young children. While children can learn problem-solving strategies from television they cannot easily transfer that knowledge because of interference from different perceptual cues. If such techniques are to be included on educational programs for young children they should be accompanied by planned activities coupled with feedback from an adult, which will help the child transfer the strategy to real-life situation.

### References

- Bandura, A., Ross, D., & Ross, S. Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, 1963, 66 (1), 3-11.
- Gagne, R. Human problem-solving: Internal and external events. In B. Kleinmuntz (Ed.), *Problem-solving: Research, method and theory*. New York: John Wiley & Sons, Inc., 1966.
- Lesser, G. *Children and television: Lessons from "Sesame Street"*. New York: Random House, 1974.



- O'Bryan, K., & Boersma, F. Eye movements, perceptual activity, and conservation development. *Journal of Experimental Child Psychology*, 1971, 12, 157-169.
- Olson, D., & Bruner, J. Learning through experience and learning through media. In D. Olson (Ed.), *Media and symbols: The forms of expression, communication, and education*. 73rd Yearbook of NSSE. Chicago: University of Chicago Press, 1974.
- Piaget, J. *Science of education and the psychology of the child*. New York: Orion Press, 1970.
- Salomon, G. Can we affect cognitive skills through visual media? An hypothesis and initial findings. *Audio-Visual Communications Review*, 1972, 20 (4), 401-422.
- Scott, M. Some observations on the cognitive ability of very young children. *Journal of Genetic Psychology*, 1973, 122, 17-25.
- Sullivan, E. Acquisition of conservation of substance through film modelling techniques. In D. Brison & E. Sullivan (Eds.), *Recent research on the acquisition of conservation of substance*. Educational Research Series #2. Toronto: The Ontario Institute for Studies in Education, 1967.
- Vurpillot, E. *The visual world of the child*. New York: International Universities Press, 1976.
- Yussen, S. Determinants of visual attention and recall in observational learning by pre-school and second graders. *Development Psychology*, 1974, 10, 93-100.

ROGER BOURGEOIS

*Université de Moncton*

and

DOYAL NELSON

*The University of Alberta*

## Young Children's Behavior in Solving Division Problems

*Sixty children, ten in each age group from three to eight years, were presented with measurement and partitive division problems involving ferrying cars across a river and then parking the ferried cars beside some houses. Thirty of these children were also presented with similar problems but, this time, involving placing animals in cages. A year later, available children were brought in again. This time, forty-four children did the problems involving the animal cages and twenty-one did the ferrying problem. All behaviors of all subjects in both sessions were recorded on video tape. Careful analysis of these children's behavior while solving the problems revealed that: (1) Physical structure of apparatus used as vehicles for problems influences the ease with which children solve the problems; (2) Partitive division appears to be a more difficult process than measurement division; (3) Requirements for one task tend to carry over to a succeeding task or tasks and influence children's choice of problem solving procedure; (4) Mathematically irrelevant aspects of a problem distract different children in different ways. A discussion of these observations is included in this report. (Dr. Bourgeois is Associate Professor of Mathematics Education at Université de Moncton and Dr. Nelson is Professor in the Department of Elementary Education at the University of Alberta.)*

The importance of developing problem-solving abilities in children is acknowledged in most texts and teachers' guides dealing with mathematics instruction. It appears, however, that methods and materials advocated for problem-solving activities are often based more on exhortation and testimonials than on empirical research (Getzels, 1964). It would seem reasonable to assume that if problem-solving is to be included in a model for mathematics learning, problem-solving situations which are appropriate must first be devised. The construction of these situations should be based on sound empirical evidence that certain situations are likely to elicit certain responses or behavior from young children.

Knowledge is lacking at present to indicate how children react to different aspects of a problem at different age levels. If some element of reality or other extraneous material is introduced into a problem, for example, how does a child behave when confronted with the situation? Does he see the mathematical aspect of the problem or is he distracted by its perceptual representation? To what degree does the apparatus or the configuration influence his response? Is there a transfer of behavior from one problem situation to another? It was in search of answers to such as these that Nelson and Sawada (1974, 1975) undertook a series of longitudinal studies.

It was the purpose of the study reported here to investigate the behavior of young children as they attempted to solve two division problems. These problems were among twelve designed for the Nelson and Sawada investigation.

### *Design*

In order to furnish valuable information on the way children perceive a task, problems were designed in such a way that they did not necessarily require verbal responses for solution. Each of the two division problems used for this study consisted of a measurement division task and a partitive division task. The two problems were, for convenience, labelled the Cargo Groups Problem and the Animal Groups Problem. A different apparatus was used for each problem.

The apparatus for the cargo groups problem, illustrated in Plate 1, consisted of a masonite base 76 cm x 106 cm x 6 mm. At each end of the board 4 mm masonite sections, approximately 76 cm x 40 cm, were glued leaving a winding strip (the river), approximately 35 cm wide, across the middle of the board.



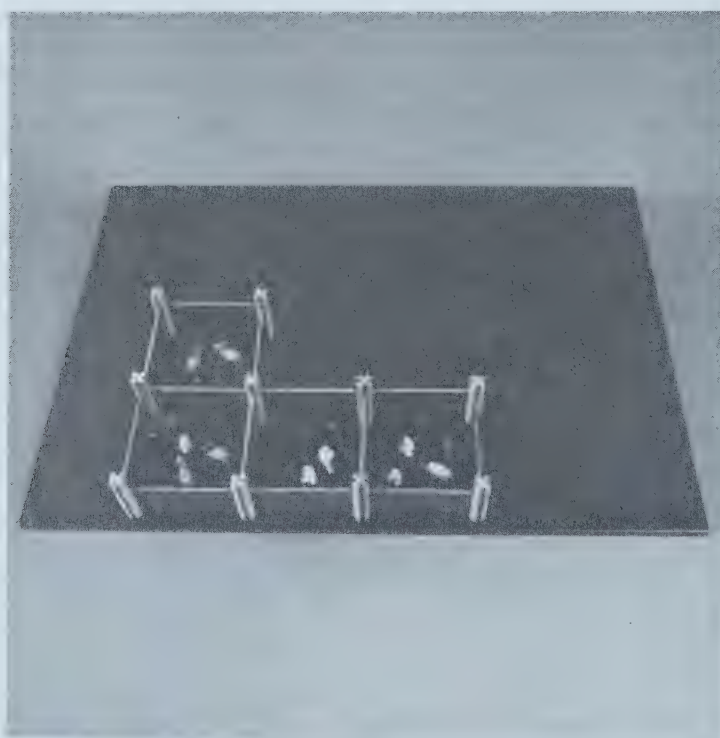


Two kidney-shaped islands in the middle of the river were made from 3 mm masonite glued to the baseboard. The river area was painted light blue and the islands were green. On each side of the river a large parking lot was painted black as was a roadway leading away from each parking lot. The remainder of the board was painted green.

There was a ferry boat of masonite with balsa wood sides, with one end convex and the other concave. The apparatus also consisted of 15 plastic toy cars of various colors. Three wooden blocks, one blue, one red and one brown were used to represent houses.

The cars were placed in the parking lot closest to the child and he was asked how many trips would be required for the ferry boat to take all the cars across to the other parking lot if the ferry boat were to take three cars at a time (measurement division). After the problem had been solved to the child's satisfaction, he was asked how many cars would be at each house if there were to be the same number at each house (partitive division).

The apparatus for the animal groups problem, illustrated in Plate 2, consisted of (i) a baseboard, (ii) posts, (iii) walls, and (iv) animals.



The baseboard, of 2 cm plywood, had dimensions 76 cm x 107 cm. Twenty-five holes formed a square array with centres 18 cm apart situated symmetrically on the board. The board was painted dark green.

Each post had four slots placed symmetrically around the circumference and the diameter was reduced at one end to fit the holes of the baseboard. The walls were made of 16.5 cm x 7.5 cm x 6 mm masonite slabs which could easily slide into the post slots. A set of twenty plastic and rubber toy animals completed the apparatus. This set consisted of four camels, four ducks, four mice, four hippopotamuses, an elephant, a moose, a horse and a lion.

The set of 20 toy animals was presented to each child who was asked to build enough cages so that there would be five animals in each (measurement division). After the task was completed, some cages were dismantled, or some were built, so that there were three, and two animals were removed. The child was now asked how many animals would be in each cage if there were the same number of animals in each cage (partitive division).

Each problem situation was presented to each child individually. The material was placed on a low table in front of the child. The interviewer sat at the child's right and adhered to a prearranged protocol. Behaviors of the subjects as they attempted the tasks were recorded on half-inch video tape. Transcription of behaviors was accomplished by devising a coding system capable of showing any action or sequence of actions. Verbalizations were recorded verbatim. Sample transcriptions of the coded forms may be found in Bourgeois (1976).

The data for the study were collected over a period of two summers. The sample for the Nelson and Sawada (1974) investigation involved 90 children, 15 at each age level three to eight. The first interviews were conducted in 1974. In this phase of the investigation sixty subjects, ten from each age group, were presented the cargo groups problem and thirty of these subjects, five from each age group, were also given the animal groups problem. This was called the cross-sectional phase of the study. Seventy-four of the original 90 subjects participated in the longitudinal phase of the study. The interviews for this phase were conducted one year later during the summer of 1975. Forty-four of the 60 subjects who had attempted the cargo groups problem in the longitudinal sampling were presented with the animal groups problem during the second interview. The cargo groups problem was also attempted by 21 of these subjects.

### *Observations*

A descriptive analysis of the children's interaction with the apparatus while attempting the problems revealed a wide range of behaviors. In this paper, discussion of these behaviors will be divided into three parts: procedures employed by subjects in solving the measurement division problems of cargo and animal groups; procedures employed in solving the partitive division problems of cargo and animal groups; and response of subjects to distracting elements in the problem situations. The discussion will be of a descriptive nature. Complete statistical data are contained in Bourgeois (1976).

### *Measurement Division*

The procedures employed by the subjects while attempting the problems ranged from highly manipulative to verbal responses. Most of the subjects up to age five were unable to solve the cargo groups problem involving measurement division. The three-year-olds for the most part could not follow the instructions. About fifty percent of the four- and five-year-olds were able to place three cars on the ferry and make the five crossings. But only one subject in this age range was able to state that it had taken five trips. This behavior persisted up to age seven. Almost all subjects placed the cars in a random order in the parking lot after taking them across on the ferry. It



appears that many subjects did not think to count the number of times they were taking cars across the river and were unable to solve the problem because the integrity of the groups was not maintained. Although several seven- and eight-year-olds chose to take the cars across the river on the ferry, about fifty percent in this age range solved the problem by either making five groups in the original parking lot or by simply stating that it would take five trips.

In general, the subjects had more success with the measurement division problem associated with the animal groups than with the cargo groups. This was especially noticeable with the four- and five-year-olds. Only two subjects solved the problem without building cages during the cross-sectional interviews while most eight- and nine-year-olds solved it in this manner in the longitudinal sampling. Very few subjects chose to solve the problem by building one cage at a time and placing five animals in each as they went along. Several subjects built more cages than were necessary. The younger subjects who did so tended to place animals in all the cages while the older subjects used only four cages. Placing an incorrect number of animals in the cages persisted up to age six.

#### *Partitive Division*

There was no appreciable difference between the proportions of success and failure in the two partitive division tasks. The number of successes increased with age. It was not until age seven, however, that the subjects appeared to have control over situations involving the partitive division process.

The procedures employed by the subjects in attempting these tasks were highly manipulative. Although the problems could have been solved by making groups on the table or by counting, none of the subjects employed these procedures during the cross-sectional interviews. Such procedures were used by eight- and nine-year-olds in the longitudinal sampling.

Quite surprisingly, the findings revealed that few of the subjects systematically distributed the objects one at a time in order around the houses or in the cages. This procedure was evident only 16 times in 155 attempts. Several of the subjects started a task by distributing three cars at each house or five animals in each cage. It appears that children using this latter procedure were influenced by the requirements of the measurement division tasks which they had attempted previously and which required them to make groups of three cars or of five animals.

After having completed the partitive division tasks, the subjects were asked if there were the same number of objects in each of the groups that they had made and also how many there were. Subjects up to age six who had failed to partition the objects equally had a tendency to state that the groups that they had made contained the same number. Most of the subjects up to age six who had partitioned the objects correctly counted the objects in each group for the purpose of verification even if they had previously stated that there were the same number in each group.

#### *Behaviors Associated with the Distractions*

When mathematical ideas are embodied in a problem-solving situation, many aspects of the situation are irrelevant to the concepts and may be



considered to be distractions. The two division problems which were designed to facilitate the study of young children's behaviors consisted of models of objects or animals with which the subjects could stimulate real world situations. In observing the procedures employed by the subjects as they attempted to solve the problems, it became apparent that many children were responding to distractions inherent in the various tasks.

The younger subjects responded more to the distractions than the older subjects but some behaviors associated with these persisted up to age eight. Several of the subjects up to age six pushed (or drove) the cars onto and off the ferry. Some subjects turned the ferry around so it would fit with the shore. Some of the younger subjects made motor sounds or pushed the ferry around an island during the crossings. A few subjects refused to place the cars on the "grass" around the houses; they preferred to park them on the road in front of the houses. Some subjects appeared to associate the colors of the cars with those of the houses.

In working with the animal groups, most subjects took particular care to assure that these were standing upright inside their cages. Most children would pick up an animal if it had fallen and carefully stand it up. Classification of the animals inside the cages was a behavior exhibited by subjects in all age groups but was more frequent among the older subjects. Some classified according to kind and others according to size. A few of the younger subjects refused to place the lion in a cage with other animals. Classifying the animals, however, did not appear to be a major factor in preventing the children from solving the problems.

Many subjects exhibited spontaneous verbalizations while attempting the various tasks. Such verbalizations appeared to be elicited mostly by the apparatus when the children were young and seemed to be more task related when they were older.

### *Discussion of the Findings*

Although the findings of the present study do not generate guidelines for designing problem-solving situations, they do indicate certain aspects of such situations that influence children's behavior. The following general statements are supported by the data.

1. *The physical structure of mathematically equivalent problems designed for young children can make some more difficult to solve than others.* The measurement division tasks for the two problems were considered to have the same mathematical structure. In each, children were required to find how many groups of a given number could be made from a given set of objects. Most of the children in the age range four years to seven years were able to form the required groups in both tasks. Yet a greater proportion of the children in the age range correctly solved the animal groups measurement division problem. It should be pointed out that the cargo groups task was presented first in the cross-sectional sampling so one might suspect that the greater ease in solving the animal groups problem could be the result of learning in the first problem. However, in the longitudinal sampling the animal groups measurement division problem was given first and it still proved to be easier.

The perceptual aspect of the animal groups task seems to have aided the

subjects in solving the problem. They could look at the cages and see how many groups they had made. This was not the case with the cargo groups problem because the integrity of the groups was lost. If children are to understand the meaning of the transformation that they make with objects, it would seem reasonable to assume that they must reflect upon the results of their actions. It is only in this way that the young child may have logical-mathematical experiences as referred to by Piaget (1970). In devising equivalent problem-solving situations for young children, care must therefore be taken to monitor the kind of experiences that they will provide.

2. *It appears that the partitive division process requires a higher level of mental operation than does the measurement division process.* A comparison of solution results for the two animal groups tasks, which had many similarities such as placing animals into cages, revealed that more subjects were able to solve the measurement division task. Furthermore, certain procedures which had been used by some of the younger subjects in the measurement division task were only employed by the eight- and nine-year-olds in the partitive division task. These procedures involved sophisticated strategies such as making groups on the table without placing animals into cages. It appears that certain procedures used by the subjects to solve the measurement division problems were not available to them until they were eight or nine years of age when the mathematical process involved was partitive division. These findings suggest that solution results should not be the sole criterion in comparing the difficulty of two mathematical processes.

3. *The requirements of one task can influence young children's choice of procedure in attempting another task when the same apparatus is involved.* More than half of the subjects used a systematic procedure in distributing the objects while attempting the partitive division tasks. However, only about 10% of the subjects distributed the objects one at a time. Several subjects started the partitive division tasks by placing three cars at each house or five animals in each cage. It could be hypothesized that some children knew that they could make groups of three or five objects in the second task of a problem because they had done so in the first task. This appears to indicate that subjects who related tasks in this manner were focusing on some aspect of the groups with which they were familiar. It would seem reasonable to assume that such interaction with the material may permit simple abstraction to take place but may not be the basis for reflective abstraction. The present investigation suggests that we need to know more about how one task may influence strategies employed in another task. Such a dependence on another task for choice of procedure may be desirable in certain tasks but undesirable in other problem-solving situations.

4. *Mathematically irrelevant aspects of a problem may distract different children in different ways.* Although the data are inconclusive, there are indications that some responses to distractions may be of an idiosyncratic nature. Some subjects responded highly to the distractions while others seemed to pay little attention to them. The color of the cars and of the houses, for example, appeared to be a very distracting element for some subjects and not for others. Some responses appeared to be related to age and other to sex. Boys had a greater tendency than the girls to simulate in minute detail real world situations in the cargo groups tasks. Boys also



appeared to be distracted by the lion in the animal groups problem while the girls paid more attention to the smaller animals.

Several authors (Dienes, 1963; Skemp, 1971; Stevenson, 1975) have pointed out the paradox we face in designing problems for young children. In an effort to make problems interesting for children we often include irrelevant details. These details help to capture the child's interest but at the same time they may distract him from the information that is central to the problem. In designing problem-solving situations, the difficulty appears to be one of maintaining interest while at the same time eliminating distractions beyond a child's capabilities. Dienes (1963) defines the problem as that of finding the laws of interaction between three variables; (1) amount of noise generated, (2) amount of noise learner is able to cut through, (3) extent of motivation of the learner.

The present investigation was not set up to differentiate those distractions that would prevent the children from solving the problems from those that would not. It does reveal, however, the complexity of the problem and the findings suggest that the role of noise in problem-solving situations needs to be monitored very closely.

This research was supported by Canada Council research grants S73-0488 and S74-1206.

#### *References*

- Bourgeois, R. D. *Young children's behavior in division problems*. Unpublished doctoral dissertation, University of Alberta, Edmonton, 1976.
- Dienes, Z. P. *An experimental study of mathematics-learning*. London: Hutchinson, 1963.
- Getzels, J. W. Creative thinking, problem solving, and instruction. In E. R. Hilgard (Ed.), *Theories of learning and instruction*. The sixty-third yearbook of the National Society of the Study of Education. Chicago: NSSE, 1964, 240-267.
- Nelson, L. D., & Sawada, D. Nature and development of problem-solving behavior in early childhood. Canada Council Project, Progress Report. Research grants S73-0488 and S74-1206. Edmonton: University of Alberta, 1974.
- Nelson, L. D., & Sawada, D. Studying problem solving behavior in young children—Some methodological considerations. *The Alberta Journal of Educational Research*, 1975, 21, 28-38.
- Piaget, J. *Genetic epistemology*. New York: Columbia University Press, 1970.
- Skemp, R. R. *The psychology of learning mathematics*. Harmondsworth, Middlesex: Penguin Book, 1971.
- Stevenson, H. Learning and cognition. In J. N. Payne (Ed.), *Mathematics learning in early childhood*. Reston, Virginia: N.C.T.M., 1975, 1-14.



J. HUNDERT

*The Sir James Whitney School  
and*

B. BUCHER

*The University of Western Ontario*

## Beginning Readers Discriminate Words by their Simplest Cue

*Previous studies show that children often discriminate words on the basis of partial graphic features such as single letters, rather than the overall word configuration. This study examined features that control discrimination among artificial words, using early-reading children. Children learned to discriminate among trigrams in lists of four. Lists were presented to each of twelve children, in a counter-balanced sequence. Each list differed in the number of usable graphic cues available to discriminate the trigrams. It was hypothesized that lists in which a single letter cue could be used to distinguish between words would be learned in fewer trials than those requiring use of multiple-letter cues, but that more errors would occur in transfer tests when the single letter was removed. The hypothesis was accepted. (Dr. Hundert is psychologist at the Sir James Whitney School and Dr. Bucher is Associate Professor of Psychology at the University of Western Ontario.)*

In learning word discrimination, a child may be taught to select from a list of printed words the correct version of a spoken word. This reading skill has been considered an important prerequisite to word pronunciation (Muehl, 1961; Zeaman & House, 1963). Reliable discrimination requires that a child learn to respond to those cues that distinguish the target word from the others presented. There may be more than one distinctive cue available for this purpose. For example the word *boy* may be discriminated from the word *red* by: any one of the individual letters, a combination of any two letters, or the entire configuration.

Which of these redundant cues the child selects may not be crucial for accurate word discrimination in a given training situation, as in the example. However, whether the child attends to simple or complex cues may affect his ability to discriminate the word in new contexts (Hundert &

Bucher, 1976). For example, after successfully discriminating the word *boy* from *red* on the basis of the initial letter, *b*, the child might be presented with the word *bat*. Confusion is likely to follow. Less confusion would be expected if the child had initially attended to a more complex cue, such as all of the letters in the word *boy*. Because of the possible relationship between redundant cue selection and transfer to new situations, it is important to study the factors determining the selection of word-cues in discrimination training.

One factor that may determine cue-selection is the position of the cue in the word. Marchbanks and Levin (1965) and Williams, Blumberg and Williams, (1970) found that young readers tend to recognize words by their first and last letters. Another possible factor is the simplicity of the word-cues. This was studied by Samuels and Jeffrey (1966). They found that a group of children who learned a word list in which words contained no common letters made more recognition errors in a transfer test than children who learned a word list with half the letters in common. In the transfer test, one letter in each word was replaced with a novel letter. These results must be regarded as inconclusive, however, because the two groups were given different numbers of transfer test trials. The former group was given eight trials while the latter was given four. It was not stated whether this difference was taken into account in the analysis of the number of errors on test trials.

The present experiment examined the hypothesis that children are more likely to attend to simple rather than complex cues when both are relevant for the discrimination. Specifically, it was hypothesized that when redundant word cues were available, beginning readers would select the simplest cue (the one containing the fewest letters). This hypothesis was tested by presenting beginning readers with lists of artificial words that differed in the availability of a single letter cue. It was expected that words in lists with an available single-letter cue would be learned in fewer trials than lists requiring use of more complex cues. Moreover, the removal of the single letter cue, once the word was discriminated, was expected to result in a greater number of errors in subsequent testing. Artificial rather than actual words were used to eliminate the possibility of familiarity with training items.

### *Method*

#### *Subjects*

The 12 eldest children from a group of 23 attending a university preschool were selected as subjects. Five were boys and seven were girls. Their ages ranged from four years, one month to four years, eleven months with a mean of four years, eight months. Written parental consent for participation in the experiment was received for each child. No formal reading instruction was given in the preschool. Subjects were seen on alternative weekdays for an average of six sessions.

#### *Design*

Subjects were trained on three types of artificial word lists. Each list contained four trigrams that varied in the complexity of the cue that could

be used to discriminate among them. "Letters" on trigrams were graphic forms. Each trigram was taught individually in two phases. In the acquisition phase, subjects learned to discriminate each trigram from the other trigram in the list by a matching-to-sample procedure where they pointed to the trigram corresponding to a word spoken by the experimenter. For each of the twelve trigrams discrimination training was followed by a transfer phase where one letter in the trigram was replaced by a new letter. Deterioration in performance, across word lists, was used to identify which cue had been used in the discrimination.

Training Material

The trigrams were constructed in two parts. The "core word" consisted of the second two letters in the trigram and the "target letter" was the initial letter. Trigram construction began with the random selection of four artificial letters from a pool of 25. These four letters were then combined in pairs to form a set of four core words. The combination pattern was ab, ca,

ACQUISITION PHASE

	DISTINCTIVE	IRREGULAR	UNIFORM
	TARGET LETTER CORE WORD		
	△○□	□○□	~○□
	⊥≠○	∫≠○	~≠○
	-○≠	□○≠	~○≠
CONTROL	△□□	∫□□	~□□

TRANSFER PHASE

	•○□	┌○□	-○□
	•≠○	┌≠○	-≠○
	•○≠	┌○≠	-○≠
CONTROL	•□□	┌□□	-□□

Figure 1. An example of Distinctive, Irregular and Uniform Word Lists during acquisition and transfer phases. The components of the target letter and "core word" are indicated and control trigram identified.



db and ac, where a, b, c, d represent four different letters. This procedure was repeated to produce three sets of core words, each with a different set of four letters. The three types of word lists were then constructed by adding a different pattern of target letters to each set of core words. An example of the three types of word lists and the artificial letters is shown in Figure 1. In the Distinctive List a different randomly selected artificial letter was added to the four core words in each of the three sets, resulting in a unique target letter for each trigram. In the Irregular List one of two artificial letters was added to each core word, so that the two trigrams shared the same target letter. The trigrams in the Uniform List were formed by adding the same artificial letter to all core words, so that all trigrams in this list shared the same target letter.

One trigram in each word list was used as a control for the tendency to focus on the first or last letter (Williams et al., 1970). As shown in Figure 1, these trigrams contained a unique middle letter. The control trigrams were used to test if subjects would attend to a simple cue regardless of its position in a word.

The set of cues that could be used to discriminate trigrams differed across word lists. These cues are shown in Table 1. The trigrams of the Distinctive List could be differentiated on the basis of the target letter alone, or four more complex cues that were also available for the trigrams of the Irregular List. Two multiple-letter cues could be used to differentiate the trigrams of the Uniform List. Other complex cues were also available; for example, trigrams could be identified by the position of particular letters. However, since these were consistent across word lists, they are not described. If, as predicted, there is selection of the simplest cues, subjects would be expected to discriminate the trigrams in the Distinctive List on the basis of their target letters.

TABLE 1  
POSSIBLE USABLE CUES IN EACH WORD LIST

Word List	Possible Cues				
	Target letter alone	Target letter & middle letter	Target letter & third letter	Core Word	Entire Trigram
Distinctive	+	+	+	+	+
Irregular		+	+	+	+
Uniform				+	+

### *Apparatus*

Each word list was printed on 11 cm x 32 cm cards with trigrams spaced 5 cm apart. The word lists were presented on a training apparatus, to control subject-experimenter interaction. The apparatus consisted of a 75 cm x 60 cm plywood front panel with side supports. Centered on the front panel were a row of four 8 cm x 8 cm plexiglas windows, which displayed a word list card. Subjects pressed one of the windows to indicate his answer. Each window was attached to a pressure sensitive microswitch leading to a control box behind the panel. Also connected to the control box were four red

pilot lights, one under each window. By turning a selector switch on the control box to a certain position a microswitch could connect to the correct pilot light. Therefore, pressing the correct window lit the associated light.

### *Procedure*

Experimental sessions lasted 20 minutes. Subjects were removed from their preschool activities, escorted to the experimental room by a female experimenter, and seated across from the experimenter, with the training apparatus between them.

Two subjects were randomly assigned to one of six complete counter-balanced sequences of the three word lists. Trigrams in a word list were trained and tested individually in acquisition and transfer phases. Subjects were presented the trigrams in a word list before proceeding to the trigrams in the subsequent word list. As another precaution against possible position effects, a random half of the word lists were presented with the target letter occurring at the end of the trigrams. This was accomplished by inverting the card on which the word list was presented.

*Acquisition Phase.* Each session began with an acquisition phase. The four trigrams were displayed in the training apparatus. The subject was required to touch the trigram that corresponded to a pronounced label. Labels had been randomly selected without replacement from 25 names of simple objects such as "man" and "horse". For example if the word "boy" was selected, the experimenter would ask "Which word says boy?" If the response was correct, the red pilot light beneath that window would light. If an incorrect response was made the experimenter would say "no" and point to the correct trigram. Following an incorrect response, the subject was given four opportunities to respond correctly to the same trigram, using the same card. Once the trigram was correctly identified, or after four correction opportunities, the trigram was represented in a different order on a new card. Training on each trigram continued until it was correctly discriminated on four consecutive trials, or to a total of 24 trials. There was no need to impose this latter criterion.

*Transfer Phase.* The acquisition of each trigram was followed by a test of the cue used in that discrimination. This test consisted of measuring transfer when all the trigrams had their target letter replaced by a single unfamiliar letter. An example of a word list presented during transfer is shown in Figure 1.

A matching-to-sample procedure was conducted similar to that described for the previous phase. However, in this condition no feedback was given to prevent trial-to-trial learning. Four transfer trials were presented.

When the subject completed the acquisition and transfer phases for a trigram, he was given an opportunity to choose a snack such as a potato chip, peanut, raisin and so on. This was to encourage the subjects to continue the task. Immediate tangible rewards were not used because of evidence of their distracting effects (Spence, 1970).

### *Rationale for Word-Cue Analysis*

The cues used to discriminate trigrams were assessed by examining relative performance on the three transfer word lists. The decision table

TABLE 2  
IDENTIFICATION OF CUE BASED ON PERFORMANCE  
DURING TRANSFER PHASES

Alternative Outcomes during Transfer			Resultant Identification of Cue
Distinctive low, Irregular low, Uniform low			entire trigram for all word lists
Distinctive high, Irregular high, Uniform high			core word for all word lists
Distinctive low, Irregular low, Uniform high			target letter plus one other letter for distinctive and irregular list and core word for uniform list
Distinctive	Irregular	Uniform	target letter for distinctive list, target letter plus one other letter for irregular list, core word for the uniform word list

shown in Table 2 shows how different outcomes were used to infer the child's cue-selection strategy. Uniformly low performance on the three word lists would suggest attention to the entire trigram for each word list. Uniformly high performance would suggest that the core word formed the basis of the discrimination. Low performance on the Distinctive and Irregular Lists with high performance on the Uniform List would indicate the use of the target letter in combination with another letter for the first two lists and the use of the core word for the Uniform List. Finally, if the subject performed poorest on the Distinctive List followed by the Irregular List and the Uniform List, different cues were used in each of the discriminations; the target letter was selected in the Distinctive List, the target letter in combination with another letter in the Irregular List and the core word in the Uniform List. This last outcome would be predicted by the hypothesis that the subject used the simplest cue.

Results

To determine if the counterbalanced sequence used to present the three word lists affected performance during acquisition and transfer, an analysis

TABLE 3  
ANOVA SUMMARY TABLE FOR NUMBER OF TRIALS DURING  
ACQUISITION AND TRANSFER

Source	Acquisition			Transfer		
	<i>df</i>	<i>Ms</i>	<i>F</i>	<i>df</i>	<i>Ms</i>	<i>F</i>
Between Subjects	2	5.6	1.2	2	6.1	1.4
Word Lists	2	128.0	35.5*	2	136.1	32.4*
Order	2	9.1	2.5	2	10.3	2.4
Residual	2	7.5	2.1	2	8.6	2.0
Within Subject	27	3.6		27	4.2	

\* *p* < .01



of variance for a Latin Square design (Kirk, 1968) was conducted on the number of trials to criterion per trigram, during acquisition. Results are shown in Table 3. No significant effect due to the order of presentation was found. Also in Table 3, the same analysis on the number of correct responses per trigram during transfer found no significant order effects. The results were analyzed to determine if the position of the target letter affected the performance. A *t*-test for dependent measures showed no significant difference between trigrams with the target letter in the initial or final position during acquisition ( $t=1.0$ ;  $df=18$ ;  $p>.10$ ) or transfer ( $t=0.8$ ;  $df=18$ ;  $p>.10$ ) phases. In the absence of a demonstrated sequence or position effect, performance on each type of word lists during acquisition and transfer conditions were pooled, without regard to order of presentation or the position of the target-letter.

### Acquisition

Results of an ANOVA for a Latin square design on the number of trials during acquisition is presented in Table 3 and indicates an effect of word lists ( $p<.01$ ). Table 4 shows the arithmetic mean number of trials to criterion per trigram for the three types of word lists. A Sheffé test for pairwise comparisons indicated that subjects took significantly fewer trials to learn the words in the Distinctive List than either the Irregular ( $F(1, 22)=26.3$ ;  $p<.01$ ) or Uniform ( $F(1, 22)=16.3$ ;  $p<.01$ ) Lists. There was no significant difference between the performance on the Irregular and Uniform Lists ( $F(1, 22)=1.9$ ;  $p>.10$ ).

TABLE 4  
MEANS FOR TRIALS TO CRITERION DURING ACQUISITION  
AND CORRECT RESPONSES DURING TRANSFER FOR  
EXPERIMENTAL AND CONTROL TRIGRAMS

Word Lists	Acquisition	Transfer
<u>Experimental</u>		
Distinctive	5.8*	0.8*
Irregular	13.0	1.4*
Uniform	11.1	2.6*
<u>Control</u>		
Distinctive	5.3	1.6*
Irregular	7.3	2.3
Uniform	7.0	2.8

\*  $p < .01$  on Sheffé test for ANOVA

### Transfer

The transfer conditions were important for identifying the cues used during the acquisition of a word. As shown in Table 3, ANOVA for Latin square designs indicated an effect of word lists during transfer ( $p<.01$ ). The arithmetic mean number of correct trials for each of the three word lists is

presented in Table 4. A Sheffé test for pairwise comparison indicated a significant difference between the Distinctive and Uniform Lists ( $F(1, 22)=1.97$ ;  $p > .10$ ), Distinctive and Irregular Lists ( $F(1, 22)=23.3$ ;  $p < .01$ ) and Uniform and Irregular Lists ( $F(1, 22)=12.6$ ;  $p < .01$ ).

As outlined in Table 2 this result indicates that the subjects attended to the target letter in the Distinctive List, the target letter in combination with another letter in the Irregular List, and the core word in the Uniform List.

Control Trigrams

The results on the control trigrams during acquisition and transfer conditions were analyzed for by a Latin Square ANOVA and are shown in Table 5 with means for word lists in Table 4. During acquisition, effect of words lists was statistically nonsignificant. Subjects took approximately the same number of trials to discriminate a trigram for each word list. During the transfer condition, there was a significant effect of word list ( $p < .01$ ). A Sheffé test for pairwise comparisons indicated that this effect was attributable to the difference between the Distinctive and Irregular ( $F(1, 22)=12.6$ ;  $p < .01$ ) and the Distictive and the Uniform Lists ( $F(1, 22)=13.7$ ;  $p < .01$ ). There was no significant difference between the Irregular and Uniform Lists ( $F(1, 22)=1.7$ ;  $p > .10$ ).

These results indicate that the availability of simple cue in the middle of the trigram allowed the subjects to discriminate the control trigrams for each word list with equal speed. During transfer, the removal of the target letter had minimal effects on the Irregular and Uniform Lists, with greater detriment to the Distinctive Lists. The lower performance on the Distinctive List can be explained by the availability of two simple cues — the target letter and the middle letter. The subjects' cue-selection was divided between those two choices. These results suggest that the subjects searched for the simplest cue for the discrimination of a trigram regardless of its position.

Conclusion

The results of this study support the hypothesis that the preschool children in this study used the simplest cue for the discrimination of word-like stimuli. This conclusion was based on two findings. First, subjects took

TABLE 5  
ANOVA SUMMARY TABLE FOR NUMBER OF TRIALS DURING  
ACQUISITION AND TRANSFER FOR CONTROL TRIGRAMS

Source	Acquisition			Transfer		
	<i>df</i>	<i>Ms</i>	<i>F</i>	<i>df</i>	<i>Ms</i>	<i>F</i>
Between Subjects	2	10.6	1.4	2	13.5	2.1
Word Lists	2	17.2	2.1	2	54.5	8.5*
Order	2	11.5	1.4	2	14.3	2.2
Residual	2	9.3	1.1	2	13.4	2.1
Within Subject	27	8.1		27	6.4	

\*  $p < .01$

significantly fewer trials to learn the trigrams of the Distinctive List, where a single letter (the target letter) could be used to distinguish between words. The selection of the target letter rather than a more complex cue facilitated the acquisition of the list. Second, although the Distinctive Lists were acquired faster, replacing the target letters with another letter produced a high number of errors during transfer. Fewer errors occurred on the Irregular and Uniform Lists, respectively. This finding is consistent with the hypothesis that the subjects discriminated words according to their simplest cue as shown in the decision table in Table 2.

This study has direct implications for reading instruction. The results suggest that when presented with a word to discriminate from other words, children attend to the simplest usable cue. Frequently this would be a single letter. Although this strategy might be effective in the training situation, it will not aid discrimination when the child encounters similar-appearing words. The child sacrifices later generalization for speed of discrimination.

It may be more efficient to teach a child to attend to the entire configuration of the word. Hundert and Bucher (1976) trained children to point to a word when it was presented with other similar-appearing words. They showed that this training not only made it easier to learn to pronounce the words but also reduced later confusion. This result indicated that the context of word discrimination is important for reading instruction. The context determines the number of cues available to the reader and can force him to attend to complex cues which, although more difficult to learn, may produce better transfer.

We would like to thank David Batstone for his valuable suggestions on an earlier draft of this manuscript, Dr. Mary Wright and Norm Greenberg for their co-operation in conducting this experiment, and especially Debra Hundert for her assistance and support.

Requests for reprints should be sent to Joel Hundert, The Sir James Whitney School, Belleville, Ontario, Canada K8P 1B2.

#### References

- Gibson, E. J. Perceptual learning and the theory of word perception. *Cognitive Psychology*, 1971, 2, 351-358.
- Hundert, J., & Bucher, B. Effect of two word discrimination training on word pronunciation in reading. In T. A. Brigham, R. P. Hawkins, J. W. Scott, & T. McLaughlin, *Behavior analysis in education: Self-Control and reading instruction*. Dubuque: Kendall/Hunt, 1976.
- Kirk, R. E. *Experimental design procedures for the behavioral sciences*. Belmont, California: Brooks/Cole, 1968.
- Marchbanks, G., & Levin, H. Cues by which children recognize words. *Journal of Educational Psychology*, 1965, 56, 57-61.
- Muehl, S. The effects of visual discrimination pretraining with word and letter stimuli on learning to read a word list in kindergarten children. *Journal of Educational Psychology*, 1961, 52, 215-221.
- Samuels, S. J., & Jeffrey, W. E. Discriminability of words, and letter cues used in learning to read. *Journal of Educational Psychology*, 1966, 57, 337-340.
- Spence, J. T. The distracting effects of material reinforcers in the discrimination learning of lower- and middle-class children. *Child Development*, 1970, 41, 103-111.
- Williams, J. P., Blumberg, E. L., & Williams, D. J. Cues used in visual word recognition. *Journal of Educational Psychology*, 1970, 61, 310-315.
- Zeaman, D., & House, B. J. The role of extension in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw Hill, 1963.



R. F. MARCUS

and

BABETTE JENNY

*University of Maryland*

## A Naturalistic Study of Reciprocity in the Helping Behavior of Young Children

*A time-sampling method with a preset category system was used to record the psychological and task helping behavior of preschool children of diverse family backgrounds. Each child also completed a sociometric test and was rated on his relative helpfulness by two teachers.*

*Help given by children was reliably reciprocated by other children. Reciprocity was stronger when the recipient was liked by the giver and weakest when the recipient was disliked. Homan's exchange theory was used to explain why high sociometric status children received more help than they gave while their low status peers had to engage in more helping activity and received less in return. A negative relationship between sociometric status and both psychological and directed help received was found. A positive relationship was found between teacher ratings of task helping and sociometric status, but this relationship was not found for sociometric status and direct observation measures of help given. Most helping behavior occurred between children of the same sex. Consistent sex differences in favor of males over females were found for total and task help given and received as well as for psychological help given. (Dr. Marcus is Assistant Professor in the Institute for Child Study, College of Education, University of Maryland and Ms. Jenny is a graduate student in the same department.)*

One does not have to search far in order to uncover the kinds of moral teachings children hear and learn by direct tuition and other means. "One good turn deserves another" and "do unto others as you would have them do unto you" are familiar phrases which indicate a generalized norm or expectation that benefits given should at some point be repaid. Only a few laboratory experiments on various kinds of benevolent behaviors have studied the variables which are likely to govern reciprocity, while no naturalistic research studies have progressed beyond the point of support for the general statement that benevolence reaps benevolence. The present study represents a naturalistic analog (to laboratory experimentation) which

attempts to demonstrate some of the factors which govern reciprocity within a nursery school setting.

Gouldner's (1960) article on the norm of reciprocity goes far beyond the simple statement of reciprocity as a universal moral principle which governs interactions between roles, groups or institutions, for which he is often given credit. His general statement of reciprocity was framed in terms of reciprocation of "helping" behavior. Reciprocity was not seen as an all or none principle, but was seen as conditional upon some of the following factors:

[the] value of the benefit and hence the debt is in proportion to and varies with—among other things—the intensity of the recipient's need at the time the benefit was bestowed ("a friend in need . . ."), the resources of the donor ("he gave although he could ill afford it"), the motives impugned to the donor ("without thought of gain") and the nature of the constraints which perceived to exist or be absent ("he gave of his own free will"). (p. 171)

This rich source of hypotheses was further augmented by a suggestion that in some cultures friendship obligates one to reciprocate and that this reciprocity might play a major role in maintaining the stability of friendships as well as other social systems. One might expect to find reciprocity varying across different levels of relationship between children. Some of these hypotheses have been investigated, at least partially, within laboratory settings.

Goranson and Berkowitz (1966) created a situation in which college women received voluntary help with a dull task, involuntary (solicited) help, or a refusal to help from a "supervisor." Women who received voluntary help worked hardest to "repay" the favor when they were given the chance to help the supervisor. On the basis of this and similar research Schopler (1970) suggested that only when help is perceived as altruistic by the recipient, and the donor is not perceived as having acted according to his own interest (e.g., is not seen as "manipulative"), is help likely to be reciprocated.

Staub and Sherk (1970) placed fourth grade children into a somewhat more naturalistic setting in which children listened in pairs to tape-recorded stories and one of the children was given the opportunity to share candy with the other child. The number of candies shared related positively to the length of time the receiver child shared his crayon during a later drawing session. Mutual friends were found to share the crayon for a longer period of time than non-mutual friends. Although the reciprocity between candy sharing and later crayon sharing was more highly related in mutual pairs than in non-mutual pairs, neither of these correlations was significant, nor was the difference between these two correlations significant. The suggestions of reciprocity varying with the affectional quality of the relationships between children could still be viewed across a wider range of relationship, i.e. friends, non-choice and negative choice relationships in order to see whether there is a significant shift in reciprocal relationships. The present study will pursue this further.

Naturalistic observation studies have shown that, in its most general sense, the norm of reciprocity operates within preschool settings. Positive behavior on the part of children tends to be correlated with positive behavior toward that child from others. Charlesworth and Hartup (1967) found that

among 3- and 4-year-olds the more reinforcements the children gave the more reinforcements they received from others. Kohn (1966) studied the social behavior of 11 kindergarten children and recorded the number of positive and negative initiations given and received by the children. Positive initiations referred to kind things said to another, invitations to interact, etc., also included directions and assistance in a constructive manner (task helping). The findings indicated a positive and significant relationship between positive acts initiated by the child and positive acts initiated toward him. Such studies have not explored any of the intricacies and qualifications of the reciprocity norm as have the writings of Gouldner and the laboratory research.

The present investigation is an examination of the norm of reciprocity within a preschool setting focusing specifically upon helping behavior. Both psychological and task helping is considered as is the positive or negative regard in which the giver holds the receiver as these factors relate to reciprocity in helping behavior. As suggested in the above literature, the importance of voluntarily given and involuntarily given help will be examined.

### *Method*

A time-sampling technique was employed using a preset 10 category system of helping behavior. Nineteen preschool children were observed over a period of 1½ months time and, toward the last third of this period, were administered a test of sociometric status. Two classroom teachers rated each of the children of 10 bipolar scales of helpfulness derived from the direct observation system.

### *Subjects*

The subjects observed and tested in this study were 19 children between the ages of 59 and 74 months ( $\bar{X}$ =66.6 months) who were attending the University of Maryland's Center for Young Children. The children were selected for the Center in such a manner that they represented diverse family backgrounds.

### *Observers*

Three observers were trained to use the time sampling technique (including one graduate and two undergraduate students). The author and one of the students served as reliability checks throughout the study while the data used in the following analyses was that of the two observers who were not aware of the hypotheses of the study.

### *The Coding System of Helping Behavior*

The coding system represents a synthesis of helping behaviors recognized in naturalistic studies by Severy and Davis (1971) as varieties of both psychological and task helping behavior, and a system used by Hartup and Keller (1960) which dealt with psychological helping within the context of nurturance. In addition, Leiserson (1975) had collected 157 anecdotes of helping behavior over a year's time with all teachers and assistants writing down episodes they considered to be of a helping nature. The four



observers worked on developing a classification system which would encompass these anecdotes as part of the training period. The following 10-category system was arrived at:

- I. Gives Protection:
  - 1. verbally defends
  - 2. physically defends
  - 3. gives warning
- II. Gives Comfort or Reassurance:
  - 4. verbally comforts or reassures
  - 5. physically comforts or reassures
- III. 6. helps out of distress
- IV. Helps with Task Verbal:
  - 7. advises, suggests, directs
  - 8. informs of third person's need
  - 9. offers instrumental help
- V. Helps with Task-Physical:
  - 10. gives instrumental help.

Analyses in this study focus only upon "Total Help" (all categories), "Task Help" (categories 7-10), "Psychological Help" (categories 1-6), and "Directed Help."

Additional distinctions were made by the observers, including whether the behavior was "directed," that is, requested or directed by another, or "nondirected," not solicited by another. The observers also recorded whether the behavior was performed by the child or received by him, the name of the child giving or receiving the help. Finally, a very brief description of the episode was recorded at the bottom of the protocol.

The time-sampling technique entailed a recording of all helping behavior given or received by a particular child for a five-minute period of time, with a randomly ordered series of children viewed successively. Absent children were picked up on the following day. Observation was done only during free play and at the beginning of snack time since preliminary observation indicated that these provided maximal opportunities for helping. The 19 children were observed for 12 units of time, or a total of one hour per child over a span of 1½ months.

### *Reliability of the Coding System*

Following the suggestions made by Patterson, Cobb and Ray (1970), which indicated that unmonitored observers might show abrupt decline in the reliability of their observations, and since there were four different observers, special care was taken to prevent such instrument deterioration. Observers were trained for one month while observing a class other than the one on which study data were collected. All observers then observed for three hours at the beginning of the data collection period (first reliability check) and then for 1½ hours after 50% of the observations had been completed (second reliability check). Throughout the entire data collection period, observers worked in pairs and thus it was possible to obtain reliability figures for these two pairs over approximately 9½ hours of observation for each. Correlation coefficients ranged from .74 to .96, indicating an acceptable reliability level.

Validity of the Coding System

The coding system has face validity in that other authors (indicated above) have recognized the behaviors as helping and by the fact that teachers collected anecdotes which they considered to be helping behaviors which were here used to modify the coding system.

In addition to the above, teachers were asked to rate each child on 10 bipolar dimensions which were derived from the 10-category direct observation coding system. Seven-point Likert scales were constructed for each dimension so that the two classroom teachers could rate the likelihood that each child would demonstrate a particular kind of helping behavior. Correlations between overall teacher's ratings of helpfulness and direct observation were found to be  $r=.42$  ( $p < .10$ ) for "total help given,"  $r=.42$  ( $p < .10$ ) for "task help given" and  $r=.50$  ( $p < .05$ ) for "psychological help given," while the corresponding correlations for teacher 2's ratings were  $r=.65$  ( $p < .01$ ),  $r=.65$  ( $p < .01$ ) and  $r=.73$  ( $p < .01$ ).

Sociometric Status

Moore and Updegraff (1964) used a picture sociometric test in which children were asked first to identify all children on a board which contained all the pictures of the children in the class and then asked to locate the pictures of four children the child liked and four children the child didn't like. This technique was used in the present study by two female research assistants. Retest scores for the Moore study, over a period of a few weeks, indicated reliabilities of .62, .52 and .78 for three different nursery school groups when the measure was calculated as number of positive minus number of negative choices. This same measure of popularity was used in the present study. Validity of the measure was supported by negative correlations between popularity and child's dependency upon adults.

Results

Sex Differences

One-way analyses of variance were performed to examine sex differences for total help given, total help received, psychological help given, psychological help received, task help given and task help received. The findings may be seen in Table 1.

The analyses revealed that significantly greater helping behavior was

TABLE 1  
ANALYSES OF VARIANCE FOR SEX DIFFERENCES

Variable	Mean Boys	Mean Girls	F
Total Help Given	25.8	13.8	8.0 *
Total Help Received	33.4	25.9	4.8 *
Task Help Given	24.0	13.4	7.6 *
Task Help Received	31.7	25.2	4.0
Psychological Help Given	1.3	.33	6.2 *
Psychological Help Received	1.7	.67	.81

performed by boys than was performed by girls in all categories of behavior except psychological help received. Since the bulk of total helping behavior was composed of task helping, the sex difference findings were expected to be similar in both instances, and were.

The data were also subjected to a series of one way analyses of variance in order to determine whether helping took place primarily within or between sexes. The results of these analyses may be seen in Table 2. In all categories of helping behavior analyzed, helping took place primarily between children of the same sex. Although the frequency of psychological helping, for girls in particular, was too small to permit parametric analyses, one may see the same picture of same sex helping as dominant. The sum of help given and received with same sex interactions versus opposite sex interactions for boys was, respectively, 13 and 4 while for girls the corresponding figures were 4 and 1. Thus, it is quite clear that helping takes place primarily within and not between the sexes.

TABLE 2  
ANALYSES OF VARIANCE FOR HELPING BETWEEN AND WITHIN SEX

Group and Variable	Mean to Boys	Mean to Girls	F	Mean from Girls	Mean from Boys	F
Boys						
Total Help	2.5	13.2	8.25*	2.8	12.1	14.8**
Task Help	2.2	12.6	9.05**	2.7	11.6	12.8**
Girls						
Total Help	4.11	.56	12.2 **	6.3	1.2	10.7**
Task Help	4.89	.78	8.79**	6.67	1.33	8.8**

\*  $p < .05$

\*\*  $p < .01$

### *Reciprocity in Helping Behavior*

Table 3 presents the intercorrelations between varieties of help given and received and their relationships to sociometric status of age.

The central hypothesis of reciprocity in children's helping behavior is supported in a number of ways: the correlation between total help given and total help received is  $r = .80$  ( $p < .01$ ), indicating that children who give greater help tend to receive more help; help given to children correlates highly with help received from children ( $r = .89$ ,  $p < .01$ ); help given to teachers correlates positively with help received from teachers ( $r = .57$ ,  $p < .05$ ). These measures of helping include both nondirected voluntarily given or unsolicited help and directed (solicited) help, although the former are far greater in number. When directed help is considered alone, the correlation between directed help given and directed help received is  $r = .00$  (note: lower reliability on this variable may have mitigated against finding a correlation). Directed help given is similarly uncorrelated with total help received ( $r = .04$ ) and help received from children ( $r = -.16$ ). Therefore, help given is reciprocated when it is given voluntarily, but not reciprocated when it is requested.



TABLE 3  
PRODUCT-MOMENT CORRELATIONS FOR HELPING MEASURES,  
AGE AND SOCIOMETRIC STATUS

Variable	1	2	3	4	5	6	7	8	9	10	11
1. Total Help Given											
2. Total Help Received	.80**										
3. Help Given--Children	.96**	.72**									
4. Help Received--Children	.88**	.87**	.89**								
5. Help Given--Teacher	-.10	.06	-.38	-.23							
6. Help Received--Teacher	-.26	.15	-.41	-.36	.57*						
7. Directed Help Given	.00	.04	-.22	-.16	.74**	.39					
8. Directed Help Received	.25	.28	.26	.19	.11	.15	.00				
9. Sociometric Status	-.08	-.02	-.07	-.03	.00	.03	-.33	-.49*			
10. Teacher 1--Total Helping	.42	.26	.41	.36	-.04	-.24	-.20	.12	.30		
11. Teacher 2--Total Helping	.65**	.56*	.66**	.63**	-.18	-.21	-.24	.02	.34	.70**	
12. Age	.11	.12	.20	.13	-.36	-.02	-.55*	.34	-.06	.05	.24

\*  $p < .05$   
\*\*  $p < .01$

Help received from teachers was found to correlate positively with directed help given ( $r=.39$ ,  $p < .10$ ) and negatively with help given to children ( $r=-.41$ ,  $p < .10$ ). Therefore, children who are helped more by teachers are more likely to respond when requested to help teachers and other children. However, the greater the help received from teachers, the less likely they were to help other children spontaneously (when not requested). This interpretation is suggested since the greater proportion of total help given and help given to children was nondirected help in kind.

*Helping and Sociometric Status*

Although direct observation measures of total help given, help given to teachers and help given to children correlate near zero with sociometric status (respectively  $r=.08$ ,  $.00$  and  $-.07$ ), sociometric status correlates negatively with directed help received ( $r=-.49$ ,  $p < .05$ ). Children who solicit help or have help solicited for them tend to be less popular with their peers. It should be pointed out that both teacher ratings of helping correlate positively yet nonsignificantly with sociometric status ( $r=.30$  and  $r=.34$  for teachers 1 and 2 respectively). Later analysis of teacher ratings of task helping will show this trend to be of greater magnitude. Since the correlation between directed help given and age is negative ( $r=-.55$ ,  $p < .05$ ) it may be that teachers and other children feel it is necessary to solicit or prompt helping in the younger but not the older children.

The sample was then divided at the mean on sociometric status and helping behavior was subjected to a 2 (sociometric status) x 2 (help given help received) analysis of variance with repeated measurement taken across the latter variable. The findings indicated two-way interaction with high status children receiving more help than they gave while lower status children gave more help than they received ( $\bar{X}$  high status-given=10.0,  $\bar{X}$  high status-received=11.2,  $\bar{X}$  low status-given=17.6,  $\bar{X}$  low status-received=14.4,  $F=4.16$ ,  $df$  1/17,  $p < .07$ ). A Newman-Keuls comparison between the

means reveal that only the differences between the first two as compared with the third of these means are significantly different ( $p < .05$ ). When both help given and received are considered together, the lower status children showed greater helping than the higher status children ( $\bar{X}$  lower status=32.0,  $\bar{X}$  higher status=21.2,  $F=19.2$ ,  $df$  1/17,  $p<.01$ ). Thus lower status children are involved in more helping exchanges, although they appear to get less in return for their assistance than do the more popular children.

### *Reciprocity, Helping and Friendship Status*

In order to see whether reciprocity in helping behavior was dependent upon the degree of friendship between children, correlations were computed between help given and received for three groups: positive choice (selected as liked), neutral choice (not selected as positive or negative choice), negative choice (selected as disliked). The correlations for help given and received for the three groups were  $r=.87$  ( $p < .01$ ),  $r=.62$  ( $p < .01$ ) and  $r=.52$  ( $p < .05$ ) respectively. Testing the difference between the correlations using a test for independent samples, only the difference between the first and last correlations was significant ( $Z=2.16$ ,  $p < .05$  for a two tailed test). Children are more likely to have their help reciprocated from someone who is liked rather than someone who is disliked.

The above, of course does not reflect the amount of help given to others who vary in friendship status. The amount of help given to positive, neutral and negative choice children was subjected to a one-way analysis of variance. The results revealed the means to be 7.6, 3.5 and 1.5 respectively, and with  $df$  2/54,  $F=7.3$ ,  $p < .01$ . Using the Newman-Keuls procedure (Winer, p. 216) for testing differences between all pairs of means, all means were found to be significantly different from one another ( $p < .05$ ).

Although low frequencies do not permit parametric analysis, it is interesting to note that of the children who gave psychological help to other children ( $n=6$ ) or received help from other children ( $n=3$ ), all psychological help was given or received from children who were liked and none was given

TABLE 4  
PRODUCT-MOMENT CORRELATIONS FOR PSYCHOLOGICAL HELPING,  
AGE AND SOCIOMETRIC STATUS

Variable	1	2	3	4	5	6
1. Psychological Help Given						
2. Psychological Help Received	.29					
3. Sociometric Status	-.17	-.55*				
4. Age	.28	-.07	-.06			
5. Help Received--Teacher	-.27	.27	.07	-.13		
6. Teacher 1--Psychological Help	.52*	.06	.26	.17	-.34	
7. Teacher 2--Psychological Help	.80**	.01	.23	.23	-.26	.74**

\*  $p < .05$

\*\*  $p < .01$

TABLE 5  
PRODUCT-MOMENT CORRELATIONS FOR TASK HELPING,  
AGE AND SOCIOMETRIC STATUS

Variable	1	2	3	4	5	6
1. Task Help Given						
2. Task Help Received	.81**					
3. Sociometric Status	-.07	.06				
4. Age	.09	.01	-.06			
5. Help Received--Teacher	-.26	.04	.07	-.13		
6. Teacher 1--Task Help	.26	.25	.51*	.00	-.19	
7. Teacher 2--Task Help	.48*	.54*	.44	.24	-.23	.59**

\*  $p < .05$

\*\*  $p < .01$

or received from nonfriend or a negative choice child. Children respond to others in distress only if they are considered friends, particularly with this personal form of assistance.

Psychological Helping

Table 4 represents the correlation between psychological helping, age and sociometric status.

The correlation between psychological help given and received is positive but nonsignificant ( $r = .29$ ), indicating that reciprocity is not present here with responses to distress. The significant negative correlation between psychological help received and sociometric status ( $r = -.55, p < .05$ ) indicates that those children who received the most psychological help tended to be the less popular children. It might be that their helping might be somewhat less acceptable to the other children and thus are given less of a chance to reciprocate this help, or they may feel less sure of their helping ability.

Task Helping

Table 5 represents correlations between task helping, age and sociometric status.

As was expected, since total help is largely composed of task helping responses, the results ( $r = .81, p < .01$ ) regarding the correlation between help given and received are almost identical with previous results reported. What is present here, but not in earlier sections is the relationship between teacher ratings of task helping and the sociometric status of the child ( $r_{t1} = .51, p < .05$ ;  $r_{t2} = .44, p < .06$ ). Children who were rated as more helpful by their teacher tended to be the more popular with their peers. Direct observation measures of task help given do not correlate highly with sociometric status ( $r = -.07$ ). Thus there is support for the relationship between teacher ratings of task helpfulness and sociometric status.

Discussion

Reciprocity was found to be present in the helping behavior of young children when both total helping and task helping behavior was considered.



Help given to teachers also appeared to be reciprocated. Reciprocity was found in the area of psychological helping, but to a lesser degree, and was not present when the helping behavior was solicited. That solicited help was not reciprocated might be explained in several ways. These findings support the Goranson and Berkowitz (1966) results with adults in which solicited help was seen as less valuable than voluntary assistance and tended not to be reciprocated. In the present study, the children who received the most solicited and psychological help were also the least popular and younger children and thus it is also possible that the reciprocation of their helping might not have been required or desired. The question of lower reliability of this item is not seen as a severe limitation since other significant correlations were found between this variable and sociometric status (see below).

The question of the value of the help being given was also considered with regard to the sociometric status of the giver. The results indicated that higher status children received relatively more help than they gave; the reverse was found for the low status children. Low sociometric status children were engaged in greater giving and receiving of help than higher sociometric status children. These results might be explained in terms of Homan's (1961) social exchange theory which would predict that the value of one child's assistance to another would be directly related to the former's overall social status, which in turn has been found related to such variables as cooperativeness with routines, social participation, friendliness, etc. (Moore, 1967). Low status children would have to engage in more helping activity than their high status peers to receive an equivalent "return" from their efforts.

Reciprocity also occurred regardless of whether the child considered the recipients his friends, nonfriends (neutral choice), or disliked peers. The correspondence between help given and received was, however, stronger for liked than disliked children and the total amount of helping given was dependent upon the child's liking of the recipient: the greater the liking the more help the child gave. This could be a reflection of the greater amount of time a child spends with his friends, thus providing greater opportunity for reciprocation. It might also suggest that receipt of help from one who regards you positively creates an obligation to reciprocate, regardless of whether the liking is mutual, whereas this obligation to repay might not be as strongly felt toward someone who dislikes you. Finally, help given by someone who likes you might be perceived as more valuable and/or less motivated by self interest than help given by one who dislikes you.

The results showed a negative correlation between psychological and directed help received, sometimes referred to as indices of dependent behavior, and sociometric status. McCandless, Bilous and Bennett (1961) found that low status children exhibited the greatest emotional but not instrumental dependency. Moore and Updegraff (1964) found a similar negative correlation for the youngest group of children between sociometric status and a composite measure which included affection and support seeking from adults. Moore (1967) suggested that either emotional dependency may be regarded as inappropriate behavior by other children or that this behavior may simply limit the amount of time the child spends establishing

friendships with his peers. The present findings indicate that the children who receive more help from teachers also tend to give less help to peers, thus suggesting that teacher help might tend to limit the amount of time a child spends helping his peers and possibly opportunities for friendship formation.

Sex differences were found here quite consistently in favor of males over females. Severy and Davis (1971) did not analyze their naturalistic findings on helping behavior for sex differences and Hartup and Keller (1960) did not find sex differences in nurturance. Neither have laboratory studies on altruistic behavior uncovered consistent sex differences (see Maccoby & Jacklin, 1974 for a review and summary of these studies). However, the research is much more convincing with regard to sex differences in aggression in which males are found to be more aggressive than females (Maccoby & Jacklin, 1974) and this may have implications for the study of sex differences in helping behavior. Greater aggressiveness on the part of boys in naturalistic settings might place them at the scene of emotional distress, at times caused by the comforter himself, thus providing greater opportunity to respond with psychological helping. Anecdotal evidence from the present study also suggests that the boys frequently engaged in gross motor play with large wooden blocks which provided considerable opportunity for help in carrying and passing them back and forth. It did not appear that the girls had a smaller number of opportunities for task help in their activities.

Other avenues of research which are suggested by the principle of reciprocity concern both the maintenance and initiation of friendship relationships. Gouldner (1960) suggested that the beginnings of new friendships might be characterized by reciprocal trading of benefits. In this way uncertainty may be reduced and trust established. It is also suggested that once such obligations are created that there is also an accompanying obligation not to harm the giver of benefits. Reciprocity might also be present while liking for another is great, but might break down as either an antecedent or a consequence of the termination of a friendship. Such developmental investigation of the initiation and maintenance of friendships could be viewed with regard to tradeoffs in a variety of benefits (helping, compliments and encouragement, etc.) which may or may not be traded off in equivalent portions. Determination of "equivalence" of benefits is itself a concept requiring further explication. A friendship between a child of high and a child of low sociometric status might require that the benefits given by the latter be much greater; the relationship is imbalanced and requires greater sacrifices by the less powerful partner, yet might be sustained despite lack of reciprocity by the more powerful partner.

Research on the concept of reciprocity entails assumptions about the nature of kind acts toward others which are quite different than those basic to the conceptualization and study of altruism. For example, the former assumes that one does not get something for nothing, while the latter does not. It is a more materialistic and pragmatic view in which helping another in need might actually take place with the giver intending to create a debt which can be collected later. Although perhaps a less noble view of man's interpersonal behavior than concepts of altruism might imply the ex-



planatory power of such normative rules has just begun to receive systematic investigation.

The authors wish to express thanks to Dr. H. Gerthson Morgan for his support of this study, Miss Beth Neiman and Miss Arlene Coopersmith for their work in collecting the data, and the Computer Science Center of the University of Maryland for its support of computer usage time. A special thanks is offered to Dr. Marilyn Church and Mrs. Marion Leiserson of the Center for Young Children for their cooperation and support of this research.

### References

- Charlesworth, R., & Hartup, W. W. Positive social reinforcement in the nursery school peer group. *Child Development*, 1967, 38, 993-1002.
- Gouldner, A. W. The norm of reciprocity: A preliminary statement. *American Sociological Review*, 1960, 25, 161-171.
- Goranson, R. E., & Berkowitz, L. Reciprocity and responsibility reactions to prior help. *Journal of Personality and Social Psychology*, 1966, 3(2), 227-232.
- Hartup, W. W., & Keller, E. D. Nurturance in preschool children and its relation to dependency. *Child Development*, 1960, 31, 681-690.
- Kohn, M. The child as a determinant of his peers' approach to him. *Journal of Genetic Psychology*, 1966, 109, 91-100.
- Leiserson, M. Helping behaviors of young children: Identifying and categorizing ways young children help each other. Occasional paper #14, Center for Young Children, University of Maryland, College Park, Md., 1975.
- Lytton, H. Observation studies of parent-child interaction: A methodological review. *Child Development*, 1971, 42, 652-682.
- Maccoby, E., & Jacklin, C. *The psychology of sex differences*, Stanford, California: Stanford University Press, 1974.
- McCandless, B. R., Bilous, O. B., & Bennett, H. F. Peer popularity and dependence on adults in preschool-age socialization. *Child Development*, 1961, 32, 511-518.
- Moore, S. G. Correlates of peer acceptance in nursery school children. In W. Hartup & N. L. Smothergill (Eds.), *The young child*. National Association for the Education of Young Children, 1967, 229-247.
- Moore, S. G., & Updegraff, R. Sociometric status of preschool children related to age, sex, nurturance-giving, and dependency. *Child Development*, 1964, 35, 519-524.
- Patterson, G. R., Cobb, J. A., & Ray, R. S. A social engineering technology for retraining aggressive boys. In H. Adams & F. Unikel (Eds.), *Georgia symposium in experimental clinical psychology*. Vol. 2. New York: Pergamon, 1970.
- Schopler, I. An attribution analysis of some determinants of reciprocating a benefit. In J. Maccoby & L. Berkowitz (Eds.), *Altruism and helping behavior*. New York: Academic Press, 1970.
- Severy, L. J., & Davis, K. E. Helping behavior among normal and retarded children. *Child Development*, 1971, 42, 1017-1031.
- Staub, E., & Sherk, L. Need for approval, children's sharing behavior, and reciprocity in sharing. *Child Development*, 1970, 41, 243-252.
- Thompson, G. G. Children's groups. In P. H. Mussen (Ed.), *Handbook of research methods in child development*. New York: John Wiley and Sons, Inc., 1960, 821-850.
- Winer, B.J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.



R. J. WILLIAMS

*Cumberland Valley School System*

and

BOB ALGOZZINE

*University of Florida*

## Differential Attitudes Toward Mainstreaming: An Investigation

*The issue of mainstreaming exceptional children has received considerable professional attention. Little has been done to investigate regular teachers' attitudes toward inclusion of handicapped children in their rooms. This study attempted to assess classroom teachers' opinions about several factors relevant to mainstreaming. The results indicated that attitudes were different for different handicapping conditions. Some implications with regard to teacher training practices and success of mainstreaming are presented. (Dr. Williams is Director of Special Education, Cumberland Valley School System and Dr. Algozzine is Assistant Professor in the Department of Special Education at the University of Florida.)*

The subject of mainstreaming has become an area of considerable professional interest; it suggests that physically, mentally and emotionally handicapped youngsters be included in regular classes with extra support from special personnel as required. Cegelka and Tyler (1970) have summarized the arguments for and against the inclusion of special children in regular classrooms.

Mainstreaming is generally supported by literature which indicates that the population of some special classes is culturally and socially dominated (Dunn, 1968); that the academic advantages do not outweigh the sociological disadvantages of placement (Christopolos & Renz, 1969; Goldstein, Moss & Jordan, 1962; Hoeltke, 1966; Jones, 1972; Mercer, 1970); and that a technology exists which can facilitate special programming by regular teachers (Christopolos, 1973; Sabatino, 1972).

Opponents of mainstreaming feel that the abolition or reduction of special classes is short-sighted and unrealistic. They argue that general educators have enough difficulty with regular students and that the addition of handicapped children will only increase the problems (Turney, 1975).

Regular class teachers themselves are unsure of their abilities to deal with exceptional children and fear that they do not have the curriculum expertise necessary to provide appropriate programs (Zawadzki, 1974).

Recent federal legislation has emphasized that the special class should not be the most desired placement; in fact, "least restrictive environments" should be given primary consideration (Abeson, 1974). Such a placement provides experiences for the handicapped child that are as close to those of normal children as possible. Regardless of professional arguments, it would appear that mainstreaming in a practical sense has become a reality.

An important variable in the success of mainstreaming programs will undoubtedly be the attitudes of regular teachers who are asked to program for handicapped children. Teacher expectations and attitudes have been demonstrated to have effects on student performance in a variety of settings (Brophy & Good, 1974; Grieger, 1971; Palardy, 1969); some evidence exists with regard to teachers' expectancies for handicapped children. Foster, Ysseldyke and Reese (1975) have shown that differential attitudes were related to labels (i.e., MR, ED, LD) assigned to a child. Algozzine, Mercer and Countermine (1976) demonstrated that behaviors were differentially bothersome as a function of the "type" of child (i.e., LD or ED) thought to exhibit the behavior. Special and regular education teachers have been shown to have different tolerances for behaviors commonly exhibited by children thought to be disturbed (Algozzine, in press).

If it can be demonstrated that teachers hold differential attitudes toward mainstreaming exceptional children, then the possibility exists that those attitudes may bias the prognosis for children integrated into "low attitude" rooms. The success or failure of mainstreaming particular types of children could be a function of differential attitudes and expectancies for different handicapping conditions.

The purpose of this study was to determine teacher attitudes toward the mainstreaming of physically handicapped, socially/emotionally disturbed, educable mentally retarded and learning disabled children. It was hypothesized that there would be no differences between regular class teachers' acceptance of children as a function of handicapping conditions and that the development of educational programs for the children would be equal for all handicaps.

### *Procedure*

*Subjects.* Three hundred and sixty-three teachers from a large Intermediate Unit in central Pennsylvania were contacted to participate in the study. Each was asked to complete a questionnaire developed for this investigation; it requested their opinions about mainstreaming four types of handicapped children. The categories were defined using the current Pennsylvania State Department criteria for learning disabled (LD), socially/emotionally disturbed (SED), physically handicapped (PH), and educable mentally retarded (EMR) children. After each definition, a brief description was given which highlighted its educational characteristics. Seventy-four percent of the teachers returned completed questionnaires ( $n = 267$ ). Approximately ninety percent of the teachers were certified in regular education, thirty-two percent had Master's degrees and less than thirty



percent were teaching or had taught with handicapped children in their room.

*Definitions of categories*

1. *Learning Disabled (LD)*. Learning disabled pupils are those children who have a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in imperfect ability to listen, think, speak, read, write, spell or do mathematical calculations. Such disorders include such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. This term does not include children who have learning problems which are primarily the result of visual, hearing or motor handicaps, mental retardation, emotional disturbance, or of environmental disadvantage. The child's intellectual potential is average or above.

2. *Socially/Emotionally Disturbed (SED)*. Socially/emotionally disturbed students are those who because of mental illness or psychological stress may exhibit overt behavior deviating from aggressive destruction to morbid withdrawal. They are students of average or potentially average intelligence but are unable to function "normally" or make progress at a rate, and to an extent, commensurate with their abilities.

3. *Physically Handicapped (PH)*. School-aged physically handicapped persons are those who manifest orthopedic and/or other health impairments of sufficient magnitude to limit their classroom accommodation and educational performance.

4. *Educable Mentally Retarded (EMR)*. Educable mentally retarded students are those who are significantly impaired in their adaptive behavior as a result of subaverage general intellectual functioning which originates in the formative years of life and is associated with impairment of one or more of the following areas: learning, maturation, or social adjustment.

*Method*. The dependent measure was the subject's responses to four sets (i.e., one for each category of handicap) of two questions designed to measure their attitudes toward mainstreaming. The first asked what portion of the exceptional child's education should be offered in the regular room; responses were recorded on a three-point continuum from "all" to "none." The subjects also responded with regard to their ability to provide a meaningful educational program for the handicapped child; a four-choice continuum, from "strongly agree" to "strongly disagree," was utilized to record these attitudes.

Two separate within subjects analyses of variance were performed to determine if the teacher's attitudes were differentially influenced by the four handicapping conditions.

*Results*

The means, standard deviations and analysis of variance summary table for the subjects' "acceptance" scores for each category of handicapped children are presented in Table 1. Follow-up procedures, using the individual variances and covariances as variance estimates, indicated that the teachers would be more "accepting" of physically handicapped and learning



TABLE 1  
MEANS, STANDARD DEVIATIONS AND ANALYSIS OF VARIANCE SUMMARY  
TABLE FOR "ACCEPTANCE" SCORES FOR EACH CATEGORY  
OF HANDICAPPING CONDITION

Category	PH	LD	SED	EMR
Mean	2.2	1.9	1.7	1.6
Standard Deviation	0.7	0.5	0.7	0.6

Note: Score range = 1-3.

Source	MS	df	F
Between Subjects			
Error	0.524	256	
Within Subjects			
Categories	15.158	3	48.740*
Error	0.311	768	

\*  $p < .01$

disabled children than they would be of socially/emotionally disturbed or educable mentally retarded youngsters.

Similar results were indicated for the analysis of the teachers' responses to the programming question; the means, standard deviations and analysis of variance summary table for this analysis are presented in Table 2. The results of the follow-up analysis are presented in Table 3; teachers felt more capable of programming for LD and PH children than for those called SED or EMR.

*Conclusion*

The purpose of this study was to investigate certain attitudes of regular teachers toward the integration of PH, LD, SED, and EMR students into regular classes; it was hypothesized that teacher attitudes would be similar for the different handicapping conditions. This hypothesis was rejected; teachers felt that PH and LD youngsters should receive more services in regular rooms and that they were better able to provide programs for them than for "disturbed" or "retarded" children.

This hierarchy of "acceptability" has been supported by other studies. Tringo (1970) found less "social distance" between nonhandicapped persons and physically handicapped individuals; more between them and mentally retarded persons. Shears and Jensema (1969) suggested that disabilities with the least social stigma were the most acceptable. Algozzine, Mercer and Countermine (1976) have posited that teachers' level of tolerance is a function of the disturbingness of a child's behavior and have shown that behaviors characteristic of LD children are less disturbing than those of SED children.

TABLE 2  
MEANS, STANDARD DEVIATIONS AND ANALYSIS OF VARIANCE  
SUMMARY TABLE FOR RATED ABILITY TO PROVIDE EDUCATIONAL  
PROGRAMS FOR EXCEPTIONAL CHILDREN

Category	PH	LD	SED	EMR
Mean	2.8	2.5	2.2	2.1
Standard Deviation	0.9	0.8	0.7	0.9

Note: Score range = 1-4.

Source	MS	df	F
Between Subjects			
Error	1.352	256	
Within Subjects			
Categories	29.419	3	54.807*
Error	0.537	768	

\*  $p < .01$

TABLE 3  
SUMMARY OF MEAN DIFFERENCES IN TEACHERS' ATTITUDES ABOUT  
THEIR ABILITIES TO PROVIDE MEANINGFUL EDUCATIONAL  
PROGRAMS FOR HANDICAPPED CHILDREN

EMR	SED	LD	PH
2.1	2.2	2.5	2.8

It would appear that the success of mainstreaming programs could be affected by the attitudes of "receiving" teachers. In-service and pre-service training programs should begin to consider methods of alleviating biases held by classroom teachers. The results of Haring, Stern and Cruickshank's (1958) study which attempted to alter attitudes provides some direction for these efforts. Their results indicated that teachers' attitudes did not change as a result of "in-service" training but that those teachers who had "experience" with handicapped children had more positive feelings about them. It would seem that teacher training programs would find more success in having students "work" with exceptional children rather than "learn" about them in classes.

#### References

- Abeson, A. Movement and momentum: Government and the education of handicapped children — II. *Exceptional Children*, 1974, 41 (2), 109-115.
- Algozzine, B. The disturbing child: A matter of opinion. *Behavioral Disorders*, in press.

- Algozzine, B., Mercer, C., & Countermine, T. Labeling exceptional children: An analysis of expectations. Manuscript submitted for publication, 1976.
- Brophy, J., & Good, T. *Teacher-student relationships*. New York: Holt, Rinehart and Winston, 1974.
- Cegelka, W. J., & Tyler, J. L. The efficacy of special placement for the mentally retarded in proper perspective. *Training School Bulletin*, 1970, 67 (1), 33-68.
- Christopolos, F. Keeping exceptional children in regular classes. *Exceptional Children*, 1973, 39 (7), 569-572.
- Christopolos, F., & Renz, P. A critical examination of special education programs. *Journal of Special Education*, 1969, 3 (4), 371-379.
- Dunn, L. M. Special education for the mildly retarded—is much of it justified? *Exceptional Children*, 1968, 35 (1), 5-22.
- Foster, G., Ysseldyke, J., & Reese, J. I never would have seen it, if I hadn't believed it. *Exceptional Children*, 1975, 41, 469-473.
- Goldstein, H., Moss, J. W., & Jordan, L. *Early school development of low I.Q. children: A study of special class placement. Interim Report July 1959 to June 1961*. Urbana: University of Illinois, 1962.
- Grieger, R. M. Pygmalion revisited: A loud call for caution. *Interchange*, 1971, 2 (4), 78-91.
- Haring, N. G., Stern, G. S., & Cruickshank, W. M. *Attitudes of educators toward exceptional children*. Syracuse University Press, 1958.
- Hoeltke, G. M. Effectiveness of special class placement for educable mentally retarded children (Doctoral dissertation, University of Nebraska, 1966). *Dissertation Abstracts International*, 1967, 27, 3311. (University microfilms, No. 67-03432)
- Jones, R. L. Labels and stigma in special education. *Exceptional Children*, 1972, 38, 553-564.
- Mercer, J. R. Sociological perspectives on mild mental retardation. In H. C. Haywood (Ed.), *Socio-cultural aspects of mental retardation*. New York: Appleton-Century-Crofts, 1970.
- Palardy, J. M. What teachers believe—what children achieve. *The Elementary School Journal*, 1969, 69, 370-374.
- Sabatino, D. A. Resource rooms: The renaissance in special education. *Journal of Special Education*, 1972, 6 (4), 335-347.
- Shears, L. M., & Jensema, C. J. Social acceptability of anomalous persons. *Exceptional Children*, 1969, 36 (2), 91-96.
- Tringo, J. L. The hierarchy of preference toward disability groups. *The Journal of Special Education*, 1970, 4 (3), 295-305.
- Turney, D. Mainstream or quiet eddy. *Contemporary Education*, 1975, 46 (2), 146.
- Zawadzki, R. F. A study of what regular classroom teachers consider deterrents to teaching the educable mentally retarded child in regular classes. (Doctoral dissertation, 1973). *Dissertation Abstracts International*, 1974, 35, 292. (University microfilms, No. 74-15613, 220)



H. W. HODYSH

*The University of Alberta*

## A Critical Examination of Dewey's Method of Inquiry

*A most critical but neglected problem confronting historians of education is the clarification of their methodological assumptions. The question is examined with particular reference to the nature and purpose of John Dewey's genetic or historic method. Contrary to Dewey's position, it is argued that the educational historian's lack of "intellectual control" over data affects the interpretation of the past. Dewey's claim that the genetic and experimental methods as tools of inquiry are identical in their application is explored and, in terms of the educational historian's unique task, questioned. Educational historians would do well to reconsider Dewey's idea of the genetic method but with a recognition of its limitations. (Dr. Hodysh is Associate Professor in the Department of Educational Foundations at The University of Alberta.)*

John Dewey identifies the central question of historical inquiry to be, "Upon what grounds are some judgments about a course of past events more entitled to credence than are certain other ones?" (1938, p. 231). That others have attempted to answer this vital question (Barzun & Graff, 1957) in no way detracts from Dewey's innovative effort to see it in terms of the genetic or historic method, a pattern of inquiry that indicates "the manner or process by which anything comes into experienced existence" (Dewey, 1902, p. 109; Ratner, 1950, p. 143). There is evidence to suggest, however, that Dewey's interpretation of the genetic method is not consistent and, in some instances, at variance with the principles of inquiry he puts forth (Murphy, 1939; Reichenbach, 1939). On the one hand, he rejects the Rankean view that the educational historian intellectually reconstructs past events "as they actually happened" (Dewey, 1938, pp. 236-237). On the other hand, he argues that application of the genetic method leads to an understanding of "the exact conditions, and the only conditions which determine a state of affairs" (1902, pp. 108-109; 1916, pp. 66-67). Such an apparent dichotomy reflects an approach to the educational past that would seem to contradict the very nature and purpose of Dewey's own historical method. The problem remains unresolved and open to further investigation.

*Historical Judgments and the Method of Inquiry*

Like others before him, Dewey conceives history to be “that which happened in the past” and “the intellectual reconstruction of these happenings at a subsequent time” (1938, p. 236; Hodysh, 1974, p. 154). The distinction between an actual and reconstructed past allows the historian of education to decide what significance should be attached to data in a continuum of past, present, and future. The past, Dewey explains, “is of logical necessity the past-of-the-present, and the present is the past-of-a-future-living present.” With the historian controlling the determination of subject-matter by applying a “*direction of movement*,” history of education becomes a cumulative continuity of events “towards stated outcomes” (1938, p. 236, italics in original). Such direction enables the investigator to see the “moral progress and moral order” which the past conceals (Hodysh, 1970, p. 248). In this context, history of education serves an ethical purpose by contributing to “a systematic account of our judgments about conduct, insofar as these estimate it from the standpoint of right or wrong, good or bad” (Dewey & Tufts, 1932, pp. 3-6).

Among Dewey’s most important considerations is the recognition that “*all historical construction is necessarily selective*” for the educational past cannot be reconstructed in its entirety (1938, p. 235, italics in original). This position acknowledges the educational historian’s centrality in determining the nature and direction of events. Though data are first selected by the people “whose past is now being written,” it is the historian who decides what “weight and relevancy” ought to be assigned to those data he chooses to study (1938, p. 236).

Dewey submits that the formation of historical judgments lags behind that of physical judgments, and this he attributes to historians who have failed to make clear “the systematic conceptual structures which they employed in organizing their data to anything like the extent in which physical inquirers expose their conceptual framework” (1938, p. 233). He holds that historical inquiry is no different from any other form of inquiry. It should be approached in terms of the “most effective operation of intelligence” which he identifies as the scientific or experimental method (Dewey, 1931 p. 326; Bhattacharya, 1969). For Dewey, there is an “exact identity between what the experimental method does for our physical knowledge and what the historical method in a narrower sense may do for the spiritual region: the region of conscious values.” In generating water, for example, “we single out the precise and sole conditions which have to be fulfilled that water may present itself as an experienced fact.” He observes that “the experimental method is entitled to rank as the genetic method,” as both are concerned with the way in which anything comes into “experienced existence.” Moreover, history provides “the only available substitute for the isolation and for cumulative recombination of experiment.” The crudeness and simplicity of early periods act as “a substitute for the synthesis of experiment” (1902, pp. 109-113).

If Dewey is correct, then a genetic explanation of a particular state of affairs  $C_t$ , occurring at time  $T$  and where  $C_t$  is the result of a series of occurrences whose initial term is some state of affairs  $C_o$  that existed before  $C_t$ , involves reference to a description of events  $C_o, C_1 \dots C_i, \dots, C_k, C_k', C_k'', \dots, C_t$  (Nagel, 1961, pp. 567-568). In this way a genetic explanation



allows the educational historian to exercise "intellectual control" (Dewey, 1966, p. 228) over his data and provides him with the opportunity to make "warranted assertions" (Dewey, 1938, pp. 154-155) that tentatively confirm or disprove the hypothesis governing his study (Dewey, 1929).

Dewey's idea of "causal linkage" plays a significant part in this genetic process. It acts as a "means of instituting, in connection with determination of other similar linkages, a single unique history." Gross, qualitative occurrences which are "logically, not temporally conjoined" resolve into a definite set of "causal" interactions (Dewey, 1938, pp. 458-459). In order that an indeterminate state of affairs might culminate in a single continuous history, "there is sufficient and necessary ground for taking one event as 'effect' or consequent, and some other as antecedent or 'cause'" (Dewey, 1938, p. 459). This results in a temporal-spatial continuity of "functionally significant" historical events (1925, p. 295) which are based on a selection of particular conditions and guided by a particular "end-in-view" (1938, pp. 460-461). In brief, the historian of education establishes the precise nature of the "cause and effect" process according to his conception of what is significant.

Application of the genetic method is evident in the *Ethics* where Dewey and Tufts inquire into the nature of "group morality" in history (1932, pp. 45-65). On the basis of secondary as well as primary source materials, Dewey and Tufts describe and compare social-historical conditions in primitive societies. Reference to such educational agencies as the family and "primitive court" is made. Historical data are interconnected and then directed in a "causal" sequence. The findings are eventually applied to the solution of ethical problems in politics, education, and family life.

### *The Limits of Genetic Explanation*

It appears that Dewey's historical method rests in part on two explicit assumptions: first, that the simplification of phenomena in history is equivalent to the isolation of phenomena in an experiment; and second, that use of this method assists in the making of ethical judgments which are relevant to the present. This raises the question of whether or not Dewey's historical method satisfies his own criteria of experimental inquiry.

Insight to the question might be had by considering Dewey's distinction between actual and reconstructed events. A physical scientist, for instance, studies phenomena as they occur in their "actual" historical state or "natural" setting. In synthesizing water, he defines and quantifies those variables needed for the completion of the experiment, thereby exerting "intellectual control" over the actual phenomena at hand. This is the goal towards which Dewey's historian strives but is unable to attain. The genetic or experimental method notwithstanding, there is no way in which the historian considers educational events as they actually happen. His data have in part been affected by the conceptual frameworks and selective processes of others. His primary sources are reconstructions of phenomena that are a step further removed from a comparable kind of "actual" data used by his colleague, the physicist. To suggest that a genetic inquiry into *similar* but *unrelated* events in time will necessarily improve the effectiveness of the educational historian's warranted assertions is to disregard the limitations of the process itself. These anomalies will likely remain as



long as Dewey considers experimental inquiry to be the *sine qua non* of historical study (Piatt, 1939, pp. 109-111; Frankena, 1965, pp. 138-139).

Such problems bring into focus a distinction between inquiry in the physical sciences and history. When conducting an experiment, for example, the physical scientist's main concern is to describe and explain "actual" phenomena within the purview of experimental control. He appeals to no immediate end beyond the determination of "the exact conditions, and the only conditions" surrounding an event. The significance of his experiment to social and ethical questions, for example, is ancillary to the primary end. Dewey's educational historian, however, sees the determination of "exact conditions" as subordinate to the "direction of movement towards stated outcomes" envisioned at the start of the inquiry. Might not such an "end-in-view" lead to predetermined goals that are apart from the experimental process and prejudicial to it?

Dewey insists that in the series *Co, Cl . . . , Ci, . . . , Ck, Ck', Ck", . . . , Ct*, the "earlier terms do not develop the later ones," but rather "constitute elements in a problem which is solved by discovering a continuous process or course" (Dewey, 1902, p. 116). Is it possible, however, to establish a "cause-effect" continuity of development from *Co* to *Ct* in terms of an impartial, scientific process, especially when the historian determines the directional movement and specific outcomes as a means of realizing the "moral progress as well as the moral order" which history conceals? In Dewey's historical inquiry into the nature of "authority" and its effect upon the social order (1946, p. 170), events are examined not in order to better understand their importance to the period under study but rather to show how they relate to judgments "of ways to act, of deeds to do, of habits to form, of ends to cultivate." It becomes clear that the historian of education establishes "significant" data for a purpose which is other than that of determining the "only conditions" of a state of affairs. Moreover, it seems difficult to justify the claim that there is an "exact identity" between the genetic and experimental methods.

### *Implications for Further Research*

At the risk of oversimplification, one might distinguish two intellectual currents affecting Dewey's historical theory. The first represents an "evolutionary" perspective that envisions the course of history as a kind of "progressive" continuity of movement towards an "end-in-view." It indicates in part a biological orientation (Passmore, 1957, p. 117) and the tendency to see historical study in service to social and moral ends (Hodysh, 1970, p. 252). The second is an "experimental" framework that identifies historical inquiry with the principles of scientific method and is in keeping with his naturalistic philosophy. These positions are noted throughout Dewey's writings on the theory of history and do not appear to reflect conceptual differences between his earlier and later works. Dewey accepts no one position unequivocally but seeks to develop a unique perspective influenced by a variety of elements (Dykhuizen, 1973). The resulting inconsistencies betray a synthesis which is not fully effected.

Abraham Kaplan tells the story of the tippler searching under a street lamp for his house key which he had lost some distance away. Asked why he

did not look for it where it was dropped, he replied, "Because it's lighter here!" (Kaplan, 1964, p. 11). In some ways, Dewey's approach to the making of historical judgments is analogous in principle to the "tippler's search." It represents a "logical" perspective particularly fitted to the special problems of physics and biology. Perhaps at this time, educational historians might question the value of Dewey's synthesis and his somewhat rigid adherence to the ideals of experimental method, and explore other "logics-in-use" in order to better understand the unique problems confronting the study of the past.

#### *References*

- Barzun, J., & Graff, H. *The modern researcher*. New York: Harcourt, Brace & Company, 1957.
- Bhattacharya, N. C. The concept of "intelligence" in John Dewey's philosophy and educational theory. *Educational Theory*, 1969, 19, 185-195.
- Dewey, J. The evolutionary method as applied to morality. *The Philosophical Review*, 1902, 11, 107-124.
- Dewey, J. *Essays in experimental logic*. Chicago: The University of Chicago Press, 1916.
- Dewey, J. *Experience and nature*. Chicago: Open Court Publishing Company, 1925.
- Dewey, J. *The quest for certainty: A study of the relations of knowledge and action*. New York: G. P. Putnam & Sons, 1929.
- Dewey, J. *Philosophy and civilization*. New York: Milton, Balch and Company, 1931.
- Dewey, J. *Logic: The theory of inquiry*. New York: Henry Holt & Company, 1938.
- Dewey, J. *Problems of men*. New York: Philosophical Library, 1946.
- Dewey, J. *Democracy and education*. New York: The Free Press, 1966 (first published in 1916).
- Dewey, J., & Tufts, J. *Ethics*. New York: Henry Holt & Company, 1932.
- Dykhuisen, G. *The life and mind of John Dewey*. Carbondale: Southern Illinois University Press, 1973.
- Frankena, W. *Three historical philosophies of education*. Chicago: Scott, Foresman & Company, 1965.
- Hodysh, H. W. The aesthetic dimensions of history. *The Journal of Aesthetic Education*, 1974, 6, 153-155.
- Hodysh, H. W. Historical theory and social change in John Dewey's philosophy. *Educational Theory*, 1970, 20, 245-252.
- Kaplan, A. *The conduct of inquiry*. San Francisco: The Chandler Publishing Company, 1964.
- Murphy, A. E. Dewey's epistemology and metaphysics. In P. A. Schilpp (Ed.), *The philosophy of John Dewey*. Chicago: Northwestern University Press, 1939, 157-192.
- Nagel, E. *The structure of science*. New York: Harcourt, Brace & World Incorporated, 1961.
- Passmore, J. *A hundred years of philosophy*. London: Gerald Duckworth & Company, 1957.
- Piatt, D. Dewey's logical theory. In P. A. Schilpp (Ed.), *The philosophy of John Dewey*. Chicago: Northwestern University Press, 1939, 103-134.
- Ratner, S. Dewey's contribution to historical theory. In S. Hook (Ed.), *John Dewey: Philosopher of science and freedom*. New York: The Dial Press, 1950, 134-152.
- Reichenbach, H. Dewey's theory of science. In P. A. Schilpp (Ed.), *The philosophy of John Dewey*. Chicago: Northwestern University Press, 1939, 157-192.

L. M. BEZEAU

*Ontario Institute for Studies in Education*

## Equality of Educational Opportunity and Inequality of Per Pupil Expenditures

*This study examines the extent to which equality has been achieved in per pupil expenditures among census divisions in Ontario. This is done by relating two measures of welfare, personal per capita income and assessment per school-aged person, to per pupil expenditures at the elementary and secondary levels. An inverse relationship was found between these two variables and per pupil expenditures at both levels, indicating that wealth of the area is still an important variable in determining per pupil expenditure. The best economic predictor of per pupil expenditures was proportion of industrial and commercial assessment, though this variable is fiscally neutral because of the percentage equalizing grant plan operating in Ontario. (Dr. Bezeau is Assistant Professor in the Department of Educational Administration at The Ontario Institute for Studies in Education.)*

Equality of education has been an important social concern in Ontario for some time. Equality of opportunity can be distinguished from equality of result (Bane, 1975 p. 281) with the latter receiving increasing emphasis through various forms of special and compensatory education. The use of pupil weighting factors in the Ontario grant structure which provides extra money to Boards with special education programs and non-Anglophone students has been directed ostensibly toward greater equality of result. A previous study of the funding of elementary education in Ontario (Armstrong, 1973 p. 11) identified two interpretations of equality, the libertarian interpretation and the egalitarian interpretation. The latter advocates compensation for disadvantages as opposed to mere equality of opportunity. Although both approaches involve redistribution of funds for education, it is generally conceded that the egalitarian approach requires more redistribution than an approach that aims merely for equality of opportunity. To attempt equality of results requires the most massive redistribution. Armstrong (1973 pp. 11-12) states that Ontario is committed to an egalitarian approach without implying that it is attempting to achieve complete equality of result.



Some authors make a distinction between equality and equity. They see equality as an easily measurable mathematical concept. Complete equality among pupils can be achieved if they have equal amounts of the resources that matter. In this study, per pupil expenditure is used as a proxy for more detailed information on resource allocation to pupils. Equity requires that the distribution of resources be fair and just in some sense. This concept clearly depends on important value judgments about what constitutes justice in society. Ideas of equity derive from some general philosophical scheme and may be abstract and difficult to measure. Equality of result, using this distinction, is really a form of equity since it is based on an unequal distribution of resources to achieve a particular view of justice.

A series of surveys (Humphreys, 1972; Humphreys & Rawkins, 1974) examined inequality between urban and rural schools in Ontario. Between 1967 and 1971 many of the services previously available only in urban areas became widely available in both. These included consultants, support personnel, special classes, and reduced class size. The trend toward equality of opportunity was reversed between 1971 and 1973. During that period the situation in rural schools remained stable or worsened, while that in urban schools continued to improve. Humphreys and Rawkins (1974) attribute this to the firmer entrenchment of the special services in urban areas in the face of provincial expenditure ceilings imposed on school boards.

A more recent study (McEwan et al., 1975) commissioned by the Government of Ontario to study educational costs questions the egalitarian motives of the Province after a detailed examination of weighting factors. They conclude that:

In any consideration of weighting factors, it is difficult to escape the conclusion that the criteria for their determination have been developed, at least in part, with the objective of accommodating the higher spending levels already established by a few boards, rather than to meet the respective educational needs of all boards. (McEwan et al., 1975, p. 234)

McEwan's Committee on the Costs of Education felt that the grant provisions as a whole were anti-egalitarian in effect:

The weighting factors, for both grant and expenditure purposes, continue to produce results that are difficult to justify. The differences among boards in the amounts of their expenditures eligible for grant illustrate the situation. A principle of the grant plan is designed to arrive at an equitable rate of grant for a board, based on equalized assessment per weighted pupil. If that principle functions as it should, some considerable equity should be achieved, subject to the inadequacies of assessment practices as we have described them earlier in this Report. But then, to apply the rates of grant calculated on this basis to different amounts of the expenditures by boards distorts the equalizing effects of the applicable rates of grant. So, while there is an attempt to treat boards on an equitable basis, the effect of the application of the total grant weighting factor is to work in the opposite direction. (p. 255)

Although the Committee looked at the grant and expenditure formulae in detail and concluded that they were inequitable, they did not attempt to assess the resulting expenditure distribution by relating it to independent variables. This study filled that gap. It examined the extent to which equality has been achieved in Ontario by examining certain correlates of per pupil expenditure at the elementary and secondary levels.

### *The Context*

The school system of Ontario is divided into elementary and secondary levels on legal and financial bases. Publicly supported separate schools exist at the elementary level for the constitutionally defined religious minority, which is usually Roman Catholic. Where separate schools exist, minority families have the option of supporting and enrolling their children in either school system. At the secondary level, all families support the non-denominational public schools through property taxes, and all children educated at public expense go to public schools. As a result, a public board of education in Ontario has different sets of ratepayers at the elementary and secondary levels. Separate school representatives on the board vote on secondary school matters but not on elementary school matters. Thus a clear line must be drawn between elementary and secondary revenue and expenditures. Some public systems in Ontario have junior high schools (grades 7, 8, and 9) and in these the law requires a detailed breakdown of classes taught at each level by each teacher to determine the correct allocation of salaries. Other expenditures are allocated in a similarly careful manner. The same distinctions are made in schools which are separate at the elementary level and private at the secondary level.

Most public school districts in Ontario are coterminous with the regional municipality, county, or district used for local government purposes. The exceptions are of two types: (1) large cities or boroughs with their own boards and (2) remote areas with large numbers of very small boards, each often operating a single school. The regional municipality, county, and district are also used by Statistics Canada as census divisions for Ontario, a practice that greatly facilitated data collection for this study.

### *The Sample and Data*

The basic unit of analysis is the census division as defined in the 1971 Census of Canada, with some exceptions. Between 1971 and 1974 a number of municipal and school district consolidations took place. Data for the original units were combined to correspond to the new units. These consolidations affected only units with relatively small populations. For the purpose of this study, the most populous census division in Ontario, Metropolitan Toronto, was broken down into its six component municipalities. Metropolitan Toronto contains 27 percent of Ontario's population, making it much too large to be treated as a single case. There are seven school boards in Metropolitan Toronto, one public board for each municipality and a separate board for the entire metropolitan area. Census data were available by municipality. The final sample contained 54 cases (census divisions, etc.) and covered the entire province.

The two dependent variables, elementary and secondary per pupil expenditures for 1974, were weighted means of the corresponding per pupil expenditures for all the boards in the census division operating at the appropriate level. This information was obtained from unpublished sources provided by the Ministry of Education of Ontario and the Metropolitan Toronto Public School Board. The weights used were the number of pupils, which resulted in a mean equal to the total expenditure at each level divided by the total number of pupils at that level. A problem was posed by some



separate boards which operated in two or more census divisions. In these cases their pupils had to be arbitrarily divided between the census divisions before calculation of the weighted mean for the elementary level. In the case of Metropolitan Toronto the number of separate school pupils assumed for each municipality was made proportional to the product of the number of children in school and the number of Roman Catholics in the municipality. Outside of Metropolitan Toronto, the division was based solely on the school age population. Although these methods are arbitrary their effect on the final results is minimal since separate boards spend only seven percent less per pupil than do public elementary boards for Ontario as a whole. Also, the high-spending separate boards tend to be in the same areas as high-spending public boards, so that the public and separate per pupil expenditure figures were close for all cases in the sample.

The population figures are provincially reported ones (Ministry of Treasury, Economics and Intergovernmental Affairs, 1975) for 1974. These include total population and school age population defined as ages five to 19. The income figures were from the 1971 Census of Canada (94-709; 95-751) and are for total income per capita. This is an overall measure of economic welfare that reflects labor force participation and unemployment as well as productivity in the labor force.

Figures on urbanization and population density are also from the Census (92-709; 98-701). Urbanization is expressed as the proportion of persons living in urban areas, which are defined to include:

- (1) incorporated municipalities with a population of 1,000 or over and having the legal status of city, town or village; (2) unincorporated places of 1,000 population or over, having a population density of at least 1,000 per square mile; (3) the urbanized fringe of (1) or (2), if it has a minimum population of 1,000 and a density of at least 1,000 per square mile. [*1000 per square mile is equivalent to 386 per square kilometer*] (12-540, p. 55)

Ontario's population is 80 percent urban but the range of urbanization among census divisions is from zero to 100 percent. The correlation coefficient between urbanization and population density is 0.47, somewhat lower than might be expected. This results from census divisions that have large non-agricultural rural areas, such as the District of Algoma, which combine high urbanization with low population density. The rural agricultural census divisions have lower urbanization ratios and higher population densities. This contrasts with the usual pattern of high urbanization and high population density.

Assessment figures (Ministry of Treasury, 1975) are for 1974 and include all taxable assessment regardless of how it is distributed between the public and separate boards. The figures on the proportion of assessment that is industrial and commercial as opposed to residential and farm is from the same source.

### *Results*

Table 1, which shows the intercorrelations among a selected set of independent variables, reveals some important relationships. The very high correlation between income per capita and urbanization dictated the elimination of one of them, urbanization, from the regressions. A surprising



TABLE 1  
PEARSON PRODUCT MOMENT CORRELATION COEFFICIENTS  
BETWEEN INDEPENDENT AND DEPENDENT VARIABLES

Variable	EXS	YPR	ASPS	PACIN	PD	LPD	UR
EXE	0.87**	0.46**	0.34*	0.57**	0.66**	0.29	0.55**
EXS		0.33*	0.37*	0.45**	0.62**	0.26	0.40*
YPC			0.54**	0.28	0.58**	0.79**	0.81**
ASPS				-0.09	0.60**	0.54**	0.26
PACIN					0.22	-0.01	0.51**
PD						0.71**	0.47**
LPD							0.61**

\* Significance  $\leq .01$   
\*\* Significance  $\leq .001$

Key:

- EXE = Elementary Per Pupil Expenditure in Dollars
- EXS = Secondary Per Pupil Expenditure in Dollars
- YPC = Personal Per Capita Income in Dollars
- ASPS = Assessment in Kilodollars Per Person 5 to 19 years of Age
- PACIN = Industrial and Commercial Proportion of Assessment
- PD = Population Density Per Square Kilometer
- LPD = Logarithm of Population Density
- UR = Proportion of Population in Urban Areas

nonrelationship is that between equalized assessment per school-aged person and the percentage of assessment that is industrial and commercial (-0.09). An examination of the residuals revealed the explanation, two types of municipalities which do not follow the high industry—high assessment pattern. The first of these are the resort areas where summer homes boost assessment without affecting resident population. These rural areas typically have little industry. The second type includes relatively poor northern mining and industrial areas which have low assessment even though a large proportion of it is industrial.

The two dependent variables, elementary and secondary per pupil expenditure, have a joint Pearson product moment correlation coefficient of 0.87, indicating about 75 percent common variance. The standard deviation of elementary per pupil expenditure was 10.4 percent of its mean as compared to a figure of 7.0 percent for the secondary level. Elementary per pupil expenditures are less equally distributed than secondary per pupil expenditures.

The best predictor of both elementary and secondary per pupil expenditure as seen in Table 1 is population density. Additional insight was gained by examining the scattergrams of the natural logarithm of population density on the horizontal axis and per pupil expenditures on the vertical axis. Population density could not be plotted without a transformation because of extreme skewness. The plots exhibited a V-shaped pattern showing that the low spending boards have population densities of 11.0 to 36.6, and that the higher spending boards have population densities between 0.14 and 11.0 and between 36.6 and 73.37. The joint distribution of the

untransformed variable is heteroscedastic with high values of per pupil expenditure being much more predictable than low values. The same conclusions hold for secondary per pupil expenditures.

The economic variable that had the strongest relationship with the dependent variables (Table 1) was "industrial and commercial proportion of assessment." At both levels this was stronger than personal income and assessment per school-aged person. The interpretation of this is not obvious. The proportion of assessment that is industrial and commercial is irrelevant in determining the relationship between the mill rate and per pupil expenditures. Ontario's percentage equalizing grant formula provides equal expenditure per weighted pupil for equal mill rates up to the grant ceiling, independently of other variables. The explanation may be that for some reason persons in industrialized areas are more willing to tax themselves to support education or that taxpayers like to see local industry support education even if it means placing a heavier burden on themselves.

The main purpose of this paper is to examine the relationship between per pupil expenditure and the wealth of an area. Per capita personal income and assessment per school age person are the two independent variables that may be considered measures of the economic well-being of an area. An egalitarian approach to financing ought to compensate areas of lower well-being by providing for higher per pupil expenditures. As Table 2 shows, this is not the case. An increase in annual personal income of 100 dollars is associated with an increase in per pupil expenditures of 7.64 dollars and 6.37 dollars at the elementary and secondary levels respectively. An increase in assessment per school age person of 1,000 dollars is associated with an increase in per pupil expenditure of 2.12 dollars and 2.63 dollars respectively at the elementary and secondary levels. Thus, if these measures of well-being have any general validity, Ontario has achieved neither equality of opportunity nor equality of result.

Additional insight can be gained by examining the results in Table 3 of multiple regressions involving the above two variables and "the proportion of industrial and commercial assessment" as independent variables. The

TABLE 2  
REGRESSION RESULTS FOR PER PUPIL EXPENDITURE  
AND MEASURES OF ECONOMIC WELL-BEING

Variables		Additive Constant	Regression Coefficients		Correlation Coefficient
Dependent	Independent		B	$\beta$	
EXE	YPC	574.95	0.0764	0.457	0.46
EXE	ASPS	719.81	2.1213	0.341	0.34
EXS	YPC	1167.56	0.0637	0.330	0.33
EXS	ASPS	1262.55	2.6314	0.366	0.37

Key:

EXE = Elementary Per Pupil Expenditure in Dollars  
EXS = Secondary Per Pupil Expenditure in Dollars  
YPC = Personal Per Capita Income in Dollars  
ASPS = Assessment in Kilodollars Per Person 5 to 19 Years

TABLE 3  
MULTIPLE REGRESSIONS OF ELEMENTARY AND SECONDARY  
PER PUPIL EXPENDITURE WITH A SELECTED SET OF  
INDEPENDENT VARIABLES

Independent Variable	Per Pupil Expenditure in Dollars				
	Multiple R	R <sup>2</sup>	R <sup>2</sup> Change	B	β
<u>Elementary</u>					
Personal Per Capita Income in Dollars	0.46	0.21	0.21	0.02039	0.12196
Assessment in Kilodollars Per Person 5 to 19 Years	0.47	0.22	0.01	2.01684	0.32429
Industrial and Commercial Proportion of Assessment	0.70	0.48	0.26	418.58399	0.56064
Constant Term	--	--	--	535.45250	--
<u>Secondary</u>					
Personal Per Capita Income in Dollars	0.33	0.11	0.11	-0.00947	-0.04903
Assessment in Kilodollars Per Person 5 to 19 Years	0.40	0.16	0.05	3.13807	0.43666
Industrial and Commercial Proportion of Assessment	0.61	0.37	0.21	432.27045	0.50105
Constant Term	--	--	--	1137.03737	--

elementary school results show that assessment per school age person has little effect beyond that of per capita income. But the proportion of commercial and industrial assessment has a strong independent variance contribution. At the secondary level, assessment has a stronger effect but here again, the proportion of industrial and commercial assessment has a very strong and independent effect on the dependent variable.

*Conclusions*

The results of comparing two measures of the economic welfare of census divisions with per pupil expenditures support the conclusions of The Committee on the Costs of Education; expenditure patterns in Ontario appear to favour neither equality of opportunity nor equality of result. The wealthier census divisions as measured by either personal per capita income or by assessment per school age person have higher per pupil expenditures at both the elementary and secondary levels. This study confirms the resource advantage enjoyed by urban areas as reported in another context by Humphreys and Rawkins (1974). The pattern reported here is undoubtedly the prevailing one in North American provinces and states. In spite of legislation, court decisions, and apparently favourable public sentiment, equality of opportunity has proven to be an elusive goal. No doubt, considerable progress has been made in Ontario and elsewhere in the past decade or two, but the old problems still exist. Equality of result which would seem to require an even more radical redistribution of revenue is still



farther from achievement. Indeed, equality of result may be simply impossible given the tremendous range of individual variation among pupils.

As indicated by the quotations from The Committee on the Costs of Education, the culprit in Ontario may be the formula used to distribute provincial funds to school boards, and especially the system of pupil weighting which is used to describe need. This is an area that needs considerably more research, both from the points of view of equality of opportunity and that of resource allocation.

#### *References*

- Armstrong, A. Ontario elementary education expenditure patterns 1970-1973. Unpublished master's thesis, The University of Toronto, 1973.
- Bane, M. J. Economic justice: Controversies and policies. In Levine and Bane (Eds.), *The inequality controversy: Schooling and distributive justice*. New York: Basic Books, 1975.
- Humphreys, E. H. Inequality and rural schools: Results of surveys in 1967 and 1969. *The Alberta Journal of Educational Research*, 1972, 18 (2), 111-123.
- Humphreys, E. H., & Rawkins, V. J. Equity and Ontario's public elementary schools. (Educational Planning Occasional Paper No. 73/74-8) Toronto: The Ontario Institute for Studies in Education, June, 1974.
- McEwan, T. A. et al. Financing education in elementary and secondary schools. (Interim Report Number Seven) Toronto: Committee on the Costs of Education, Government of Ontario, 1975.
- Ministry of Treasury, Economics and Intergovernmental Affairs. *1974 municipal financial information*. Toronto: Government of Ontario, 1975.
- Statistics Canada. *Dictionary of 1971 census terms*. Ottawa: 1972. 12-540.
- Statistics Canada. *Population: Urban and rural distributions*. Ottawa: 1973, 92-709, 1.1-9.
- Statistics Canada. *Labor force and individual income: Total income for Canada, provinces, and census divisions*. Ottawa: 1974. 94-709, 3.1-9.
- Statistics Canada. *Population and housing characteristics by census tracts: Toronto*. Ottawa: 1974. 95-751, CT-21B.
- Statistics Canada. *Geography: Land areas and densities of statistical units*. Ottawa: 1973. 98-701, SG-1.

J. C. POWELL

*University of Windsor*

D. J. COTTRELL

*Workmen's Compensation Board*

and

MARGARET LEVER

*Board of Education, Windsor*

## Schools I Would Like to See: An Opinion Survey Instrument With Interesting Possibilities

*Following a suggestion by Carlton (1974), and using as a starting point the goals for education reported by Spears (1973), a 24-item opinion survey was developed. The survey requires the selection of the four most important and the four least important goals from the list. A complex scoring procedure derives seven largely independent scales from these eight responses. Although the list of goals is not exhaustive, it has been found that this set of goals and the underlying construct upon which they were developed form a necessary and sufficient set to satisfactorily characterize the general approach to education among most populations. Reliabilities are high enough for some use in individual attitude assessment on most of the scales. Among other findings, a four-level Maslow-like hierarchy was detected to form one scale. This scale did not validate as a construct with Shostrom's POI (1966). Some examples of the use of the opinionnaire are reported and several suggestions for possible research use are made. (Dr. Powell is Associate Professor of Educational Psychology at the University of Windsor, Mr. Cottrell is a vocational rehabilitation counsellor for the Ontario Workmen's Compensation Board in Toronto, and Ms. Lever is Computer Liaison Officer for the Board of Education in Windsor.)*

Carlton (1974) reviewed the literature on schooling and reported seven different images commonly held by theorists in education: (1) The Familial Model, (2) The Political Model, (3) The Economic Model, (4) The Religious Model, (5) The Collegial Model, (6) The Custodial Model, and (7) The Community Model.

This paper reports the development of an instrument designed to empirically validate these seven models and an additional one, (8) The Humanizing Model, as a necessary and largely sufficient set of approaches to education to accommodate most people's views of what purposes education should serve. These eight models are first described in summary form and the procedure used to develop the instrument and its general characteristics is discussed. Following this, the results of initial attempts at determining the validity and the reliability are presented and a brief overview of some of the research outcomes is given. The paper concludes with a synopsis of the effectiveness of the instrument to date and some suggestions for research and other uses for which it might be employed.

### *Description of Models*

Consideration is given to these models in the order Carlton (1974) discusses them. The description of each model, as follows, is abridged from Carlton's discussion and is then summarized in terms of goal expectations.

1. *The Familial Model.* In this model, "custom and consensus are reinforced through consistency of socialization across home and school: change may be valued or even encouraged, but only along lines which amplify rather than challenge the established order" (p. 62). A person supporting this model would be expected to favour an emphasis upon extrinsic rather than intrinsic motivation, to encourage social as contrasted with individual responsibility, and to encourage learning for use in society.

2. *The Political Model.* "The political school emphasized its linkages with all levels of government, . . . [e]ffective political socialization, . . . [a]ppropriately edited curricula . . . [to] enhance the adaptive and reproductive functions of the school, . . . [without] challenge [to] the rightness of established order" (p. 64). The person who supports this political point of view would be expected to favour emphasis upon extrinsic motivation in terms of one's responsibility to the state, the development of an informed citizenry, and the suppression of personal needs and goals in favour of collective needs and goals.

3. *The Economic Model.* In this model, "economic norms govern the enterprise . . . schools are drawn into competition, . . . differentiation of the training [and the products] offered" (p. 66). In a phrase, students are sensitized to their market value. The person who favours this model for schooling would expect the school to focus upon the competitive development of vocational competence, the pursuit of economic rewards, and personal development through self-initiative.

4. *The Religious Model.* A school in this model "has behind [it] the full weight of spiritual and ecclesiastical authority. . . . Linkage to the church is . . . immediate and direct. . . . [Such a school is] protective of the status quo, . . . [and] inner directed" (pp. 67, 68). The person who favours such an approach to schooling would place a strong emphasis upon moral and ethical or spiritual values in an attempt to heighten self-criticism for the development of inner directed (intrinsic) motivation. An emphasis upon traditional learning and upon acquiescence to persons in authority would also be evident.

5. *The Collegial Model.* "The collegial school is client-centred, but . . .



ensures the regeneration of the elite. . . . The authority system . . . is hierarchial by expertise . . . [and] the education of recruits to a discipline is regarded as . . . vital” (pp. 69, 70). Thus the adherent to this approach to schooling would be expected to favour interpersonal competition, academic excellence, and individual initiative tempered by logic and supported by evidence.

6. *The Custodial Model.* “Entry to a custodial school is [prescribed] . . . teacher-student relationships take on much of the character of staff-inmate interactions . . . reinforced psychologically by depersonalized tactics . . . behavioral control [and] humiliation . . . [and] the continued denial of personal competency through external regulation. [This assumed incompetency] functions to legitimate the need for custody” (p. 71). The supporter of the custodial approach, viewing the learner as incompetent and education as therapeutic, would be expected to favour strong discipline, focus upon basic intellectual skills, and an emphasis upon the responsibility of the individual for his or her own behaviour.

7. *The Community Model.* In this model, “ [E]ducational aims are not set by a single group or organization; rather, each pursues his own goals in conflict or cooperation with others similarly self-directed . . . objectives are . . . set by those who pursue them” (pp. 72, 73). A person who would support this model would be expected to favour the development of collaborative action for the common good.

When examining these seven models in the light of the descriptions Carlton gave and the secondary comments provided here, it became apparent that a three-dimensional system could be used to organize all seven models into a single system. Table 1 gives this organization.

This system organizes these models in the following dimensions:

I. *Approach to Motivation* which refers to whether the model sees the learner’s rewards for progress as coming from within (Intrinsic) or from some external source (Extrinsic).

TABLE 1  
A THREE-DIMENSIONAL SYSTEM ACCOMMODATING  
CARLTON’S MODELS

School Model	Dimensions		
	I Approach to Motivation	II Approach to Information	III Approach to Authority
1. Familial	Extrinsic	Doing	Consensus
2. Political	Extrinsic	Knowing	Consensus
3. Economic	Extrinsic	Doing	Individual
4. Religious	Intrinsic	Knowing	Consensus
5. Collegial	Intrinsic	Knowing	Individual
6. Custodial	Extrinsic	Knowing	Individual
7. Community	Intrinsic	Doing	Consensus
8. Humanizing (?)	Intrinsic	Doing	Individual

II. *Approach to Information* which refers to whether or not information is taught for its intrinsic value (Knowing) or its utilitarian value (Doing).

III. *Approach to Authority* which refers to whether or not the ultimate authority for the right to act is vested in society, tradition or other external consensus-based foci (Consensus), or within the personal responsibility of the individual to operate in society with the authority of personal competence and conviction backed by reason and evidence (Individual).

Once these seven models were established within this framework, it became evident that an eighth model was needed to complete the system. The missing model was intrinsically oriented in the Motivation dimension, action oriented in the Information dimension, and personal responsibility oriented in the Authority dimension. Recent literature concerning attempts to “humanize” the schools would seem to reflect these characteristics, thus the suggestion for the name of the eighth model:

8. *The Humanizing Model*. The person who supports this model would be expected to favour educational goals which advocate the development of individual self-competitive initiative with a focus upon the individual’s responsibility to deliver a high level of personal service to others.

Viewed from this perspective, it seems reasonable to suggest that Carlton was rather hard on the Custodial Model, and that each of these models have a functional value in the appropriate circumstances. Criticism of any model may arise philosophically from any person holding an alternative view, or pragmatically from the inappropriate application of that model.

At this point, Table 1 may be regarded as a hypothetical construct with respect to the goals for education. It captures some, but not all, of the major conflicting points of view held concerning the purposes of schooling. The authors, intrigued and challenged by this construct, decided to test its viability empirically.

TABLE 2  
LIST OF GOALS FROM KAPPAN SURVEY<sup>a</sup>

1. Develop skills of reading, writing, speaking and listening.
2. Develop pride in work and a feeling of self-worth.
3. Develop good character and self-respect.
4. Develop a desire for learning now and in the future.
5. Learn to respect and get along with people with whom we live and work.
6. Learn how to examine and use information.
7. Gain a general education.
8. Learn how to be a good citizen.
9. Learn about and try to understand the changes that take place in the world
10. Understand and practice democratic ideas and ideals.
11. Learn how to respect and get along with people who think, dress and act differently.
12. Understand and practice the skills of family living.
13. Gain information needed to make job selections.
14. Learn how to be a good manager of money, property and resources.
15. Practice and understand the ideas of health and safety.
16. Develop skill to enter a specific field of work.
17. Learn how to use leisure time.
18. Appreciate culture and beauty in the world.

<sup>a</sup> Spears (1973)

*Design of the Opinion Survey*

There are many lists of educational goals in the literature, but Spears' (1973) report of the goals questionnaire listing 18 such objectives, circulated among 1,020 Kappans was of particular interest because so many of them could be fitted into the same models as those presented in Table 1.

To complete a list of goals which had one representative for each model in each of the three dimensions would require 24 goals, so that Spears' list was nearly enough. Table 2 gives the lead statements of each of the 18 goals in Spears' list. The statements circulated were actually more elaborate than this.

For present purposes, consensus among several people as to the placement of each of these goals was used, and six more goals, designed to fill the empty spaces with statements appropriate to the model and dimension, were added.

The initial validation attempt revealed that the three R's goal (No. 1) was rated first by nearly everyone, reducing the discriminating power of all items. Two other goals were correlated with others in the list implying that these two were not identifying unique components. These three were replaced so that the complete list ultimately contained 15 from Spears, and

TABLE 3  
GOALS LIST FROM FINAL VERSION OF  
"SCHOOLS I WOULD LIKE TO SEE"

1. Develop a capacity for action based on knowledge.	13. Develop ability to trust our own judgement.
2. Develop a pride in our work and in our status in life.	14. Develop skills for innovation.
3. Develop a sense of group solidarity.	15. Learn to manage money and other resources effectively.
4. Learn to respect and get along with others.	16. Understand and practice ideas of health and safety.
5. Develop skills for continued learning.	17. Learn to cooperate with people holding viewpoints which differ from ours.
6. Learn how to manage information.	18. Appreciate culture and beauty.
7. Develop skills of self-exploration.	19. Develop self-initiative.
8. Learn the rewards of good citizenship.	20. Develop an abiding set of spiritual and moral values.
9. Learn to use leisure time effectively.	21. Recognize and accommodate the differing needs of others.
10. Understand the changes in the world.	22. Accept our own short-comings.
11. Understand and practice democratic principles.	23. Learn to accept the authority of people in positions of responsibility.
12. Understand and practice family living skills.	24. Learn about and understand the interdependence of people.

© Canadian copyright number 260567-111, June 1975, by J. C. Powell, Windsor, Ontario, Canada. Reproduced with permission.



nine new statements of goals. Several of the statements were slightly reworded to accommodate these construct placements.

Having everyone rank order all the goals (as done in the Spears report) proved to be time-consuming and of low test-retest reliability. Alternatively it appeared that support for one pole of any dimension suggested the rejection of the opposite, implying a forced choice procedure. If these goals represented a continuum, then the highest and lowest goals chosen from this set might be sufficiently representative of a person's value to give an adequate picture of the views of education held by that person.

The choice of four high and four low goals from the total set proved to be the most reliable (in a test-retest sense). Three, five, and six alternatives were also tried but all had lower reliabilities, six being the lowest. The rank order with the high or low choices proved to be less reliable than the simpler "most important" and "least important" designations. Four at each end out of 24 represented greater removal than one standard deviation from the mean.

The final version of the instrument was composed of the 24 goals for education listed in Table 3. The title given the list was, "Schools I Would Like to See" and the respondent chose eight of these: four most important and four least important as seen by him. The instrument took less than 10 minutes to administer and mark sense sheets were easily used.

#### *Validation Procedures*

Several forms of validity were considered, the first being content validity. All were goals for education, but there was the possibility that two or more of the respondents saw them as so similar that they did not discriminate the entire model. To resolve this issue, a principal components analysis procedure was used ( $N = 242$ ) to determine the independence of each goal. None of the goals correlated significantly and the lowest loading for any item upon its own component was satisfactorily high ( $+ .63$ ); most were  $.80$  or higher.

The second form considered was concurrent validity. To obtain this validity, the 15 goals which overlapped with Spears' list were rank ordered by assigning  $+1$  for "most,"  $+5$  for middle (unselected) and  $+9$  for "least," on the basis of the average rank assigned. This rank order was then correlated with the rank order found by Spears. This correlation was very high ( $r = .95$ ). This time the sample size was larger ( $N = 450$ ). The single major departure was the goal relating to "democratic principles" (Table 3, No. 11). This was seventh out of the 15 from Spears' results, and eleventh out of 15 here. Since Spears reported a survey in the United States and the sample used here was of university students in Canada, this difference may well be a systematic difference arising from the characteristics of the two populations. It was gratifying to find such a simple instrument would be so sensitive to differential effects.

The third validity considered was construct validity. The first concern was to determine whether or not the three assumed dimensions would be found among these data. The cluster analysis procedure in the BMD package (which is related to Guttman's 1944 *Simplex*) was used. The diameter algorithm in this procedure produced three eight-member clusters

which, except for the misplacement of six alternatives, recapitulated these three dimensions precisely. These six items were reclassified on the basis of Table 1 by exchanging them to fit the hypothesized pattern more exactly. One pair was exchanged between two dimensions and the other two pairs were exchanged within each of two of the dimensions, exchanging polarity. The other two algorithms (maximizing and minimizing) were also applied to these data and two new scales were produced by these procedures.

IV. *Approach to People*. The maximizing algorithm produced two clusters of 12 members each. One contained goals mostly people- and self-oriented. The other 12 were things- and ideas-oriented. Because of this polarity, this scale was called Approach to People (People-Things). A few items in this scale seemed oddly out of place. For instance, the "groups solidity" item (Table 3, No. 3) appeared in the "Things" pole. Attempting to change them had a devastating effect upon reliability. An interesting research topic would be to explore the implications of these peculiarities.

V. *Approach to Development*. The minimizing algorithm produced three clusters of approximately six items, with each of the others standing by itself.

Examination of the item components of these clusters suggested a four-part scale reminiscent of Maslow's (1968) hierarchy:

*Level 1. Relating world to self.* These goals suggest a tendency to an egocentric, self-protective approach to education. (Example: Develop skills of self-exploration.)

*Level 2. Relating self to world.* These goals suggest a tendency for an outward looking view of education constrained by rather short temporal and special frameworks. (Example: Learn the rewards of good citizenship.)

*Level 3. Competent functioning.* These goals suggest a tendency toward an inward self-critical approach leading to an outward stress on the competitive development of competence. (Example: Develop skills for continued learning.)

*Level 4. Mediative outreach.* These goals seem to suggest a tendency toward attempts of the individual, on a self-competitive basis, to extend oneself beyond oneself. (Example: Develop skills for innovation.)

The poles seemed to be Self-protection and Self-development.

Scale IV added a fourth area in which controversy has occurred (the people-things dichotomy) and scale V added a possibility of studying Maslow's theories in an educational context.

VI. *School Model Supported*. These two new scales proved helpful in an additional approach to construct validity which was attempted. In this procedure, the three "most important" choices representative of each school model were used as a starting point. The most common fourth choice found to go with at least two of these three was taken as expected choice number four. By using consensus among a group of raters as to which three alternatives a supporter of each model would be most likely to reject, and the most common rejection not in this list, four "least important" choices for each model were also assigned. In this manner a unique "ideal" profile was obtained for each model. These "ideal" profiles were scored as though they were actual selections produced by someone who had responded to the instrument.



TABLE 4  
COMPARISON BETWEEN EXPECTED AND OBSERVED  
"IDEAL" VALUES OF THE CONSTRUCT

School Model	Expected Dimension Sign			Observed Dimension Sign			Proportion of occurrence (N = 1423)
	I	II	III	I	II	III	
1. Familial	-	+	-	-2	+2	-2	.03
2. Political	-	-	-	-3	-2	-3	.04
3. Economic	-	+	+	-2	+1	+3	.12
4. Religious	+	-	-	+1	-2	-1	.07
5. Collegial	+	-	+	+3	-1	+2	.13
6. Custodial	-	-	+	-2	-1	+3	.03
7. Community	+	+	-	+2	0	-2	.08
8. Humanizing	+	+	+	+3	+2	+3	.20
9. Indeterminate	?	?	?	?	?	?	.30

1.00

With the scores for the first three dimensions in hand, a comparison could be made from these scores with the expected signs of the scores from Table 1 in the original construct. For scaling purposes, Extrinsic, Knowing, and Consensus poles were taken as negative.

In Table 4 only the Knowing-Doing dimension for the Community Model does not agree with the construct. Twenty-three out of 24 (96 percent) fitted the construct. The original hypothetical system seems to be strongly supported by this procedure, as these eight categories seem to form a necessary set for describing any population.

To establish whether or not this construct containing these eight models forms a sufficient set, 1,423 high school students were scored using a profile matching procedure which involved classifying the respondent as supporting a particular school model. Three or more choices had to agree with the "ideal." Preference was given to the "most important" set in an ambiguous case. If the ambiguity remained or if less than three could be matched, the person was included in a ninth "undetermined" group. For the classification system to be sufficient, 40 percent or less should be in the undetermined group. This percentage is derived from the fact that a three-dimensional set with three levels rather than two would have 27 categories rather than eight. Thus, a larger number of undetermined respondents would suggest that the system was insufficient to describe any population. Table 4 also gives the validating sample results. There is 30 percent in the "undetermined" category. In no case within several studies has there been more than 40 percent undetermined. Hence, the construct would seem to be supported as "sufficient."

This profile matching supplied a sixth scale (this time a nominal scale) for the instrument which identifies which of the eight models the respondent would seem to support.



VII. *Approach to Education*. A seventh scale was also developed. The procedure in this case was to plot the “ideal” profiles on their scores for scales IV and V on a grid. The principal diagonal (through quadrants I and III) was drawn and the distance from the origin of the orthogonal projection of these points to the diagonal was determined. This gave a scale value for each school model on a dimension which has a *people-things* relationship in one direction and a *self-development, self-protection* relationship in the other. Scale VII, because of the concepts it related, was classified as an “Approach to Education” dimension with poles “Openness” and “Closedness.” The dimension gives one possible operational definition of Openness in education, making possible an alternative empirical approach to yet another controversial area in education. The various school models can be defined on the scale in accordance with their “degree of openness” with a *people to self-development* orientation being “open” and *things to self-protection* orientation being “closed” (liberated-alienated, if you prefer). To get a score for respondents, the orthogonal projection of their plotted position on scales IV and V is averaged with the value of the scale score for the “ideal” school model with which they were identified. Those in the undetermined groups were given an “ideal” score of zero. This latter value does not affect the sign of the projected value, only its magnitude. A reasonable approach for people who are “on the fence,” so to speak.

Table 5 shows the “ideal” scores for each model in scales IV, V, and VII, half the models closed and half open, with scale V being the most influential. It is not surprising to find Custodial and Religious models Closed nor Economic, Familial, Humanizing and Collegial models open. To find the Community models closed seems strange in the light of all the favourable publicity it has received. To find the Political model the most closed of all, though unexpected, would seem to be both reasonable and enlightening. Since the typical teacher tends to be nonpolitical, the finding

TABLE 5  
“IDEAL” SCORES FOR THE EIGHT SCHOOL MODELS  
ON SCALES IV, V, AND VII<sup>a</sup>

School Model (VI)	Scales		
	IV	V	VII
1. Political	-4	-3	-5
2. Custodial	-4	-2	-4
3. Community	-2	-1	-2
4. Religious	0	-2	-1
5. Economic	0	+1	+1
6. Familial	+6	+1	+5
7. Humanizing	+4	+3	(+5) +6 <sup>b</sup>
8. Collegial	+2	+9	+8

<sup>a</sup> School models are arranged in order of their values on scale VII.

<sup>b</sup> Humanizing was given a +6 value to distinguish it from the Familial Model in the scoring program. This choice was made because all scale values were positive for the Humanizing Model and not for the Familial Model.

that a sample of 242 teacher trainees tended to be open and contained no adherents of the Political model among them is not surprising.

Thus, in a construct validity basis, the instrument appears to make sense. Further evidence in support of the construct validity of particular scales will be reported later.

Predictive validities have yet to be examined although studies are in progress in this latter area.

### *Reliability*

With respect to the reliability of the test, once again several approaches were used. First, the principal components analysis accounted for .84 of the variance which suggests a high internal consistency. Second, test-retest intercorrelations were obtained, collected from several administrations with different scoring procedures. The highest reliabilities were found under the following conditions:

1. Scales I, II, and III were rearranged to fit the *a posteriori* cluster analysis pattern.
2. Scale IV was also most reliable as found in the cluster analysis. Exchanges of alternatives which seemed logically to be out of place (such as Group Solidarity among the Things Oriented) destroyed the correlation.
3. Scaling Level 1, as -2 and Level 4 as +2 in dimension V improved the reliability of that dimension.
4. Choosing the four most and four least important gave the highest test-retest reliabilities. Either more or fewer alternatives in these two sets proved less reliable. Also, treating *most* and *least* as single categories proved more reliable than using rank ordering within these categories.
5. Since scale VI is a nominal scale, the proportions of equivalent classification from one administration to another was used as a variance estimate, the square root being the reliability estimate.
6. An alternative form involving easier language for administration to younger children is under development; alternate form reliabilities will be available at a later date.

The results of the best reliability estimates ( $N = 16$ ) are given in Table 6.

It should be remembered that the first three dimensions accommodate only eight alternatives each. Since the respondent chooses only eight, these scales are based upon an average of  $2\frac{2}{3}$  responses in each scale. To get the overall stability for these three scales, the Spearman Brown formula was applied to the geometric mean of the three reliabilities. This procedure is the origin of the .76 composite value. As indicated in Table 6, the reliability of scale VI is not a correlation. Sixty-nine percent (11 out of 16) respondents reclassified into the same model on the second administration. Since the square root of .69 is .83, this figure is reported.

Since scale VII is a composite scale of all other scales, its highest possible value would be the product of the three primary components used to make it the composite of scales I, II, and III along with scale IV and V. This expectation is met exactly ( $.76 \times .81 \times .78 = .48$ ).

Table 7 gives the intercorrelations between the seven scales ( $N = 24$ ). Only three correlations are significantly different from zero, two of them in

TABLE 6  
TEST-RETEST RELIABILITIES (*N* = 16)

Dimension		Composites
I	.55	.76
II	.53	
III	.48	
IV	.81	
V	.78	.48 (Expected)
VI	.83 <sup>a</sup>	
VII	.48	

<sup>a</sup> Not a correlation

TABLE 7  
INTERCORRELATION AMONG THE SEVEN SCALES IN INSTRUMENT

Scales	Scales					
Scales	II	III	IV	V	VI	VII
I	.13	.06	.02	-.07	<u>.43</u> <sup>a</sup>	.34
II		-.19	.27	.13	-.15	.02
III			-.34	-.35	.11	-.28
IV				-.11	.18	<u>.54</u>
V					-.33	.24
VI						<u>.73</u>

<sup>a</sup> Underlined correlations are significantly different from zero.

scale VI. The correlation of .73 between scale VI and scale VII can be explained in that 50 percent of the value assigned to scale VI is used to produce scale VII. Aside from this, scale IV also contributes significantly to scale VII, suggesting that the people-things orientation is the most potent determinant of Openness or Closedness. This is a pattern which makes sense.

The other significant correlation, scale I (Motivation) with scale VI (School Model) suggests that Closed School Models tend to rely upon Extrinsic motivation and Open School Models tend to encourage Intrinsic motivation. This observation is not only sensible, but adds fuel to the controversy over the use of Behavior Modification procedures. These procedures stress Extrinsic motivation which is probably entirely appropriate in a therapeutic (Custodial) setting and by the transcendence of the State over the individual Political Setting. The relationship of Extrinsic motivation to the Community and Religious settings is uncertain. Both favour Intrinsic motivation in the model. Also, both the Economic and



Familial models appear to support Extrinsic motivation (money and family name respectively). Only 18 percent of the variance is involved. One cannot overgeneralize, but the observation is intriguing and could bear further investigation.

Because of the dependency relationship between scales VI and VII, these two scales should be interpreted separately.

Details of the scoring procedure are available in Powell (1975). The administration of this opinionnaire is quick and easy. Scoring, however, is too complex for convenient hand scoring in any quantity. A computer scoring program is available.

### *Application Studies*

The discussion which follows considers the validity and reliability data first. Then the results of the application of scale VI (the eight models) to several different groups are discussed. Third, the implications of using combinations of the six other scales are also considered.

Basically, four different studies are considered. These involve (i) about 900 students using a test-retest administration in a college of education and subgroupings thereof; (ii) about 70 teachers and 1,400 students in a secondary school involving a single administration; (iii) all staff members in a residential school for disturbed adolescent girls, and (iv) a group of 24 adults who took the Personal Orientation Inventory (POI) (Shostrom, 1966).

### *Implications of Validity and Reliability*

The validity studies are primarily of a construct nature and the general construct seems to be supported. The reliabilities are high for a procedure which requires only eight responses and, in general, are adequate. All scales except V and VI essentially make bipolar distinctions and the reliabilities are sufficiently high to use this instrument on an individual basis for such distinctions. Scale V makes distinctions among four levels and has a reliability sufficiently high that group distinctions can be made in this many levels. Individual distinctions are less certain. Scale VI is not reliable enough to be certain of group distinctions for 9 categories (including an undetermined category). However, the method used for obtaining reliability estimates for nominal scales is not very powerful. Conclusions drawn from using scale VI should be considered in the light of their reasonableness as well as their stability. Scale VI should be used independently of other scales because of the fact that it significantly correlates with scales I and VII and may contaminate the results if used with these two for interpretive purposes.

### *Interpretations from Scale VI*

The validation studies make it clear that these eight school models form a necessary and sufficient set of expectation typologies to classify most populations. Population differences are interesting; geography teachers are least definable by model, and religious studies teachers are most definable in one high school. The majority of the latter adhere to the Religious model. An overall reduction in the proportion of respondents not defined may occur with increasing teaching experience. High school students show no such phenomenon with age. Female students choose goals which place them most

frequently in the Humanizing Model. Male students show no single tendency toward preference.

Different occupational groups tend to display differential model preferences. Social workers, for instance, seem to favour the Collegial model, whereas child care workers in the same residential school for girls favour, in general, the Custodial model (Bolus, 1975). The girls in the school expressed more favourable relationships with the child care workers than with the social workers. The residents were not given the SCHOOLS survey for political reasons. These findings in aggregate attest to the validity of the Carlton-Powell system and give some hints as to its usefulness.

### *Interpretations Based on Other Scales*

In general, systematic differential effects can be found in the educational process when this instrument is used. These effects may be used as indirect evidence for the effectiveness of specific programmes.

When used on a longitudinal basis, changes in attitudes follow logically from systematic characteristics of courses and/or programs. For instance, if a teacher training program gives considerable theory but fails to supply a large amount of direct practice in applying this theory, then student teachers become significantly more self-protective (scale V) and closed (scale VII) across their training year. Alternatively, to get a significant Openness shift, the combination of theory, practice in simulation, practice in the field plus specific training in observation and testing techniques is needed. Omitting any of these four components from the program maintains the trend, but makes the change not statistically significant. This combination of four elements is not sufficient by itself to make these student teachers significantly more self-development oriented (scale V) although the trend is evident (Powell, 1976). What is needed in the latter case to produce such a change is not yet known. The trend from the procedures reported for scale V suggests that considerable exploratory involvement in the field may be necessary. Other scales change significantly or show trends that also support the apparent necessity of a multidimensional approach to education in order to produce significant effects.

Where cross-sectional studies are concerned, many interesting findings emerge. Students in high school generally move (on scale V) from self-protection (about 2/3 of grade 9) to self-development orientation (about 2/3 of grade 12). Since both theory and practice are needed for self-developmental orientation to occur, it would seem that most grade 8 students are too young to be expected to make future-oriented decisions. If such decisions are to be made by grade 8 pupils, as is Ontario practice, considerable training and practice in making future-oriented decisions will have to be added to the elementary school program. It is not known for certain how much of this self-development shift is a maturational or a drop-out phenomenon. There are no grade 9 students in level 4 on scale V and about 20 percent of the grade 12 students are at this level. The exact opposite occurs with level 1. However, the decline in level 1 could be from attenuation. A longitudinal study would be necessary to determine these relationships.

Students who opt for work in open areas involving independent study and/or interactive approaches tend to be more open (scale VII) than other



students. These differences become significant for those who are involved in both independent study and interactive study programs. It is not known if this openness is a product of the choice factor or of a learning factor. A longitudinal study is necessary in this case, as well. Many other interesting findings have emerged from the cross-sectional study of the secondary school and will be reported separately.

Although there are similarities between the apparent pattern of scale V (Self-protection; Self-development) and Maslow's hierarchy of needs, this scale seems to be measuring something quite different from the self-actualization measure claimed for Shostrom's (1966) POI. More than half of the POI scales correlate positively with scale III (Approach to Authority). Since the positive end of the scale is toward independent action, clearly many scales measure this aspect of attitude. Scale VI (School Model) correlates curvilinearly with the Self-Actualizing Values scale on the POI. The POI, therefore, did not seem to discriminate between self-initiators who are manipulative in intent (Political, Custodial) and those who are facilitative in intent (Collegial, Humanizing). Scale V did not relate to any part of the POI in a manner which supported the construct validity of either instrument. Thus, work with this present instrument seems to support Maslow's general concept of a hierarchy of needs but not Shostrom's operationalization of the concept. This hierarchy seems to be related to the environmental orientation of the learner so far as the results of SCHOOLS is concerned and would seem to be more complex than the POI measures. Other interesting negative and curvilinear relationships were found (Werner & Powell, 1976) which support this last conclusion or, alternatively, challenge the construct validity of SCHOOLS or the precise definition of Maslow's construct depending upon how the data is interpreted. More research in this area is needed to clarify these issues.

### *Conclusions*

The development of an instrument, SCHOOLS I WOULD LIKE TO SEE, to test the validity of a construct derived from Carlton's (1974) seven Popular Images of the Schools would seem to have been accomplished with a reasonable degree of success. The construct seems to be supported to a point where it is probably valid and involves a necessary and sufficient set of categories (with one addition) for the reasonable classification of most populations. The overall construct may give some interesting insights into how the schools might be functioning.

Several attempts at using the instrument produced interesting results, one being that significant differential effects among educational programs have been found in all cases where these effects could be expected. These effects also occurred in directions which make sense with respect to the nature and the structure of these programs. Since differential effects in educational research have rarely been found consistently, the instrument would seem to have promise.

Attempts to confirm the appearance of Maslow's hierarchy in scale V failed to support the construct validity of SCHOOLS or the construct validity of the POI or of the manner in which Maslow describes his hierarchy, whichever point of view one wishes to take. Scale V, however,



seems to represent a genuine phenomenon, developmentally and instructionally related. Further research is required to establish the parameters operating in this area. One possible explanation for these negative results arises from the observation that scale V would seem to measure a process, whereas Maslow's self-actualized individual and the POI derived from it would seem to be defining a state of existence. Perhaps states and processes are independent.

The observations give rise to other questions. For instance, do students who do best in a particular class, subject, or program hold views of what a school should be doing similar to those of the teachers involved? Are there particular points of view held by student which predispose them to conflict within the school; to absenteeism; to voluntary or involuntary withdrawal? Are there particular approaches to education which are more suitable for students with particular attitudes? Are there teaching or counselling procedures which can be used to change the attitudes of students from the present attitudes held by disruptive students to attitudes which are more productive toward harmony and toward quality achievement?

Are some of these eight approaches to education more useful than others, or does each model have a specific (definable) place in the educative process? Are there particular teacher attitudes which are counter-productive to effective education in a general sense? In a particular setting sense?

Concerning the school and the community: is the school giving the community the kind of education seen as necessary or desirable by the community being served? Are the conflicts between the school and the community, at least in part, a product of differing perceptions of the role the school should play? Should educators assess and try to meet community attitudes, should they try to influence community opinion toward more appropriate objectives for the schools, or should they do both?

Do board members, school officials, etc., hold views of education which are appropriate to the optimal development of the system as a whole? What kinds of human and other resources are needed to operationalize any or all of these models which seem to be appropriate within a school system? How might conflicting objectives be resolved in such a way that needs are met in some rational manner?

All of these questions and many others may be open to exploration by using the SCHOOLS opinion survey. These data are easily and quickly collected, but a computer program is required to score the questionnaire if large numbers of the opinionnaires are collected.

It is hoped that this instrument continues to sustain its initial promise as a useful tool for the exploration of education problems. Inquiries concerning research use of the instrument are welcome.

#### *References*

- Bolus, C. R. A proposed model treatment program for the Province of Ontario based upon a study of Maryvale—A treatment centre for emotionally disturbed adolescent girls. Unpublished doctoral dissertation, Wayne State University, 1975.
- Carlton, R. A. Popular images of the school. *Alberta Journal of Educational Research*, 1974, 20 (1), 59-74.

- Guttman, L. A basis for scaling qualitative data. *American Sociological Review*, 1944, 9, 139-150.
- Maslow, A. H. *Toward a psychology of being*. 2nd ed. New York: Van Nostrand, 1968.
- Powell, J. C. *Manual for Schools I Would Like to See*. Windsor: Privately published, 1975.
- Powell, J. C. Validation of a teacher evaluation instrument. Paper presented to the annual convention of the Ontario Educational Research Council, December, 1976.
- Shostrom, E. L. *Personal Orientation Inventory (POI)*. San Diego: Educational and Industrial Testing Service, 1966.
- Spears, H. Kappans ponder the goals of education. *Phi Delta Kappan*, 1973, 55 (1), 29-32.
- Werner, R. E., & Powell, J. C. The Personal Orientation Inventory compared with an independent source: Characteristics each scale measures. Unpublished Research Report. University of Windsor, 1976.

## BOOK REVIEW

TEACHING ELEMENTARY SCHOOL MATHEMATICS THROUGH MOTOR LEARNING. *By Robert B. Ashlock and James H. Humphrey.* Springfield, Ill.: Charles C. Thomas Publishers, 1976.

With current trends and emphases in elementary school mathematics focusing on active learning procedures involving use of concrete aids and games, a book that explicitly dwells on motor learning is timely if not overdue. Although there have been many book-length publications dealing with activity learning in mathematics (e.g., Biggs & MacLean, 1969; Kidd, Myers & Cilley, 1970; Dumas, 1971; Reys & Post, 1973; Schall, 1976), none have focused on the motor learning component involved. Thus it was with considerable expectation that I began examining the text by Ashlock and Humphreys—I had hopes of finally reading some in-depth presentation of the physiological, perceptual, and cognitive bases for activity learning as related to mathematics. On this score, I was profoundly disappointed; the purpose of the authors was not to relate activity learning to motor skills, or to perceptual-motor integration, or to cognitive (information processing) abilities, but rather to help the teacher to “do” motor learning of mathematics in the classroom. Although there is a chapter on the nature of learning through motor activity (Chapter 2) and a chapter on the research on learning mathematics through motor learning (Chapter 4), very little insight or information is provided on understanding how and why motor learning of mathematics works.

Specifically, the book does not provide information on how motor learning activities provide for the integration of physical, physiological, perceptual, and cognitive processes. Without a conceptual model for this integration, the teacher is left with little basis for thinking about what she is doing.

In regard to the research basis of motor learning of mathematics, the authors devote a chapter to their own research and that done by their graduate students. Five of the studies involved comparisons of the treatment A versus treatment B type of research. Three of the studies resulted in NSD findings; one study obtained significant differences but involved motor learning versus no instruction, hardly a valid comparison; the last study obtained significant differences in favour of motor learning at “a moderately high level of probability.” Given these kinds of results, the authors very aptly conclude the chapter on research by stating



However, it is interesting to note that no study has shown a significant difference in favor of traditional procedures over the motor activity learning procedure. (p. 66)

Despite making the above statement, the authors go on to conclude that they have now provided "some objective data to support" the motor learning activity approach to mathematics. This conclusion is debatable to say the most. If the book doesn't provide the conceptual basis for understanding motor learning of mathematics, and if the research cited does no more than suggest that the motor learning approach will do no worse than traditional approaches, what does the book provide? It does provide a useful and well presented set of motor learning activities that classroom teachers should find relevant. In addition to discussing general ways of providing mathematical experiences through motor activities, the authors devote four chapters to specific procedures under the headings:

1. Mathematics Motor Activity Stories
2. Learning about Number and Numeration Systems through Motor Activity
3. Learning about Operations of Arithmetic through Motor Activity
4. Learning about Other Areas of Mathematics through Motor Activity.

As well, they provide a listing of the activities by concept and by title. Enough detail is provided for each activity so that teachers interested in amplifying the motor activity aspect of their teaching could do so. Many of the activities are couched in the game format; thus the likelihood of student enjoyment of the activities appears to be high. Moreover, most of the activities are of the large muscle variety (involvement of the whole body as opposed to just hands). As such they would complement nicely the small muscle activities usually found in mathematics classrooms. Involvement of the whole body makes participation much more dynamic, forcing attitudinal and physical as well as cognitive processes to be fully functioning. Under such circumstances, it would be difficult for the student to behave as a spectator.

### Conclusion

The book does little to enhance current understanding of motor learning as a means for providing effective learning of mathematics. It does, however, provide a handy source of motor learning activities that show promise for learning mathematics.

D. Sawada  
*Dept. of Elementary Education*  
*The University of Alberta.*

### References

- Biggs, E. E., & MacLean, J. R. *Freedom to learn: An active learning approach to mathematics*. Don Mills, Ontario: Addison-Wesley Ltd., 1969.
- Dumas, E. *Math activities for child involvement*. Boston: Allyn & Bacon, 1971.
- Kidd, K. P., Myers, S. S., & Cilley, D. M. *The laboratory approach to mathematics*. Chicago: SRA, 1970.
- Reys, R. E., & Post, T. R. *The mathematics laboratory: Theory to practice*. Boston: Prindle, 1973.
- Schall, W. E. *Activity-oriented mathematics: Readings for elementary teachers*. Boston: Prindle, 1976.









## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centred and in capital letters, e.g., TABLE 4. The heading of the table is to be centred below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. the position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last name at the end of the manuscript under the heading, *References*.
6. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
7. In matters of style, the APA Publication Manual is considered definitive.





# ajer

THE ALBERTA JOURNAL OF  
EDUCATIONAL RESEARCH

VOLUME XXIII    NUMBER 4    DECEMBER 1977

PUBLISHED BY  
THE UNIVERSITY OF ALBERTA    ●    EDMONTON

## THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

*A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.*

.....

Published quarterly in March, June, September and December  
by the  
Faculty of Education  
of  
The University of Alberta

.....

### Faculty Publications Committee

<i>M. A. Assheton-Smith</i>	<i>W. W. Laing</i>
<i>N. C. Bhattacharya</i>	<i>R. G. Martin</i>
<i>J. Chambers</i>	<i>E. Miklos</i>
<i>W. T. Fagan</i>	<i>C. H. Preitz</i>
<i>H. W. Hodysh</i>	

*Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.*

Chairman: *E. Miklos*

Editor: *Henry W. Hodysh*

Secretary: *Mrs. A. Onishenko*

.....

AJER gratefully acknowledges support from Canada Council and AACES. AJER is indexed in the *Canadian Education Index*; appropriate articles are abstracted in *Educational Administration Abstracts*, *Psychological Abstracts*, and *Sociology of Education Abstracts*.

The subscription rate is \$8.00 per year; single copies are \$2.50 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available in bound and unbound volumes; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications to the Editor, *The Alberta Journal of Educational Research*, Faculty of Education, The University of Alberta, Edmonton, Alberta, Canada T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436



# The Alberta Journal of Educational Research

Volume XXIII, No. 4

December, 1977

## CONTENTS

Cognitive Processing and Reading Difficulties: A Framework for Research .....	245
<i>J. Cummins and J. P. Das</i>	
Improving Skill in Applying Mathematical Ideas: A Preliminary Report on the Instructional Gaming Program at Pelham Middle School in Detroit .....	257
<i>L. E. Allen and J. Ross</i>	
Some Effects of Distractions in Nonverbal Mathematical Problems .....	268
<i>J. P. Bana and D. Nelson</i>	
Logico-Mathematical and Spatial Development in Children Underachieving in Arithmetic .....	280
<i>L. Whyte</i>	
Teaching Superordinate Concepts with Simulation Games .....	298
<i>R. B. Cohen and R. H. Bradley</i>	
Comparative Views on School Discipline .....	305
<i>S. C. T. Clarke and S. Hunka</i>	
BOOK REVIEWS	
<i>Applying Behavior-Analysis Procedures with Children and Youth,</i> B. Sulzer-Azaroff and G. R. Mayer .....	317
Reviewed by D. Baine	
<i>Teaching Low-Achieving Children Reading, Spelling and Handwriting: Developing Perceptual Skills with the Graphic Symbols of Language,</i> A. M. Markoff .....	320
Reviewed by G. Malicky	

FACULTY OF EDUCATION  
*The University of Alberta*



## Editorship of AJER

The Publications Committee is pleased to announce the appointment of Dr. Henry W. Hodysh, Associate Professor in the Department of Educational Foundations, as Editor of *AJER*. Dr. Hodysh completed his doctoral studies at the University of Alberta; his special interests are the history and philosophy of education and comparative education.

We look forward to his term in office as Editor and wish him well in his new responsibilities.

Before he joined the staff of the Faculty of Education in 1967, Dr. Hodysh served as a teacher with the Edmonton Public School Board. Since his appointment he has been involved in teaching at the graduate and undergraduate levels and conducting research into the history of education. From 1972 to 1975, Dr. Hodysh served as Assistant Dean in the Faculty of Education.

The results of his research have been published in various journals including *Educational Theory*, *History of Education Quarterly*, and *Paedagogica Historica*. Our readers will recall that Dr. Hodysh has also been a contributor to *AJER*.

The Publications Committee wishes to thank Dr. Andrew K. Clark for his service as Editor and for his contributions to the continued development of this journal.

E. Miklos, Chairmen  
Faculty Publications Committee

# *Announcement*

## of the

# CENTRE FOR RESEARCH IN TEACHING

The *Centre for Research in Teaching* has been established in the Faculty of Education of the University of Alberta. The Centre, which is the first of its kind in Canada, will have linkages with at least two similar units in the United States. Dr. D. A. MacKay has been appointed Coordinator of the Centre.

The Centre has as its main purposes, for the period July, 1977, to August, 1980, the following:

1. To facilitate involvement of staff members of the Faculty of Education, as well as other scholars, in research in teaching and teacher education.
2. To disseminate research findings that have particular significance for teachers, teacher educators, and research scholars.
3. To plan and implement a program of coordinated research in teaching and teacher education that will involve staff and graduate students of the Faculty of Education.
4. To ensure continuation and further expansion of this line of research beyond the three-year "seed period" by developing and presenting proposals for major funding to various external agencies.

Anyone wishing to contact the Centre may write to the following address:

Dr. D. A. MacKay, Coordinator  
Centre for Research in Teaching  
Room 234, Education I  
The University of Alberta  
Edmonton, Alberta, Canada  
T6G 2G5

# new from Manchester

TEACHING LOCAL HISTORY W B Stephens Designed mainly for teachers and student teachers at junior, middle and secondary school level, this book discusses the case for the inclusion of local history in the curriculum. The practical difficulties are examined, and suggestions made for work both in and out of the classroom. There is much information on source material, and the use of printed and pictorial matter, as well as a substantial section on field-work and archaeology. Illustrated £5.95 net

A RAGGED SCHOOLING Growing up in the Classic Slum Robert Roberts  
'Moving and enthralling ... filled with humour, hope and human warmth. The finest and truest picture of the other side of the Edwardian era.' - Daily Telegraph  
'The autobiography of an exceptional man ... quite extraordinary.' - Observer  
'Marvellous ... Roberts watched, observed, marvelled - and remembered.' - Spectator Illustrated £3.75 net

PHILOSOPHY OF EDUCATION A SELECT BIBLIOGRAPHY Compiled by John P Powell The object of the bibliography is to provide a guide to the rapidly growing and somewhat scattered literature on the philosophy of education. The third edition of this important reference work is now available; it contains many new items and has extended its classification scheme to include many previously neglected areas of enquiry. Each item is numbered for cross-reference and there is a full author and journal index. paper £2.25 net

RESEARCH IN EDUCATION Edited by A D Edwards and J Rushton The main emphasis of the journal is on empirical research in all the recognised fields of education, together with interdisciplinary studies; and includes papers on research in education at all levels. There is also a regular book review section. The journal is published twice yearly.  
1978 subscription \$15, single copies \$8

THE TEACHER IN A CHANGING SOCIETY Edited by John D Turner and James Rushton £3.50 net

EDUCATION AND DEPRIVATION Edited by James Rushton and John D Turner £3.95 net

EDUCATION FOR THE PROFESSIONS Edited by John D Turner and James Rushton £4.50 net

MANCHESTER UNIVERSITY PRESS Oxford Road, Manchester M13 9PL



JAMES CUMMINS

and

J. P. DAS

*The University of Alberta*

## Cognitive Processing and Reading Difficulties: A Framework for Research

*The present paper outlines the potential of the simultaneous-successive model of cognitive processing as a framework for understanding and ultimately remediating reading difficulties. Recent studies suggest that among groups which might be expected to experience some difficulty in reading, successive processing is highly related to reading performance. However, at more advanced levels of reading, simultaneous processing is equally or more important for the development of fluent reading skills. The implications of these findings for research on reading disabilities are discussed. (Dr. Cummins is Research Associate at the Centre for the Study of Mental Retardation and Dr. Das is the Director of the Centre, Faculty of Education, The University of Alberta.)*

Although the phenomenon of dyslexia or specific reading disability is still very inadequately understood, in recent years neuropsychological research has begun to delineate the brain-behaviour relationships which underlie certain types of reading difficulties. In this paper we shall synthesize recent research findings on the cognitive processing and linguistic abilities of reading disabled (RD) children within the framework of Luria's neuropsychological theory of simultaneous and successive processes. We shall first briefly outline Luria's theory and examine some recent studies which have related simultaneous and successive processing to the reading achievement of various groups of children. Then we shall attempt to show the potential of Luria's theory as a framework within which to investigate the cognitive processing of reading disabled children.

Luria (1966, 1970, 1973) has proposed a neuropsychological model which divides the human brain into three functional systems or blocks. The first block is responsible for maintaining wakefulness and arousal and consists of the upper brain stem, the reticular formation and the older parts of the limbic cortex and hippocampus. The second block includes the posterior cortex

consisting of parietal, occipital and fronto-temporal lobes and is responsible for the input, recoding and storage of information. Both simultaneous and successive processing occur in this second block. The third block of the brain consists of the prefrontal lobes and is responsible for the construction and execution of plans or programs. Through extensive connections with the other two blocks, it regulates and controls purposeful conscious action.

Simultaneous and successive processing were suggested by Luria on the basis of his observations of different types of cortical lesions and their behavioural correlates. Simultaneous processing involves the integration of the individual stimuli arriving at the brain into simultaneous and primarily spatial groups and is linked to the occipito-parietal area of the cortex. This type of processing is required in the formation of any holistic gestalt or in the discovery of the relationships among two or more objects. Although simultaneous synthesis has spatial overtones it is involved in several types of linguistic processing tasks. For example, understanding of logico-grammatical relations (e.g., "father's brother"), comparative constructions (e.g., "taller than"), and spatial prepositional constructions (e.g., above, below, etc.) all depend, according to Luria, on simultaneous processing.

Successive processing, in contrast, is linked to the fronto-temporal regions of the cortex and integrates individual stimuli into temporally organized successive series. As examples of successive processes Luria (1966) gives the automatization of skilled movements, rote memory, and narrative speech.

In a large number of studies (reviewed by Das, Kirby, & Jarman, 1975) it has been found that factor analyses of different types of reasoning and memory tasks can be parsimoniously interpreted in terms of simultaneous and successive processing. Simultaneous, successive and speed factors have emerged in studies involving many different clinical and cultural groups, e.g., retarded, learning disabled, native Indian, East Indian, as well as normal children. Thus, different cultural and clinical groups employ similar strategies of processing information. However, differences may exist between groups in preference for and proficiency in a particular strategy.

### *Linguistic Functioning and Simultaneous-Successive Processing*

Luria's description of simultaneous and successive processes suggests that both modes of processing are involved in the individual's linguistic functioning. Because of the sequential nature of language and speech, successive processes are clearly important. However, as noted above, simultaneous processes are involved in understanding complex conceptual-linguistic relationships. We have recently investigated the relationship between simultaneous and successive processing and performance on a variety of linguistic tasks at the grade 3 level (Cummins & Das, 1977). Briefly, it was found that successive processing was significantly related to children's ability to analyze the linguistic form of propositions, specifically propositions which involved ambiguities in surface structure and deep structure. Simultaneous processing, on the other hand, was significantly related to performance on verbal reasoning tasks. In a previous study (Cummins, 1973), we have found that a syllogistic reasoning task involving comparative constructions and a task involving the discovery of semantic relationships between words both loaded on a simultaneous factor in high school students.



These findings are consistent with predictions derived from Luria's theory and suggest that simultaneous and successive processes are differentially important for different aspects of linguistic functioning.

However, the role of simultaneous and successive processing in linguistic functioning may vary between groups and at different developmental levels. It is possible for example, that successive processing of the linguistic input may be a prerequisite for deeper levels of semantic analysis involving simultaneous processes. Thus, in the reading process, successive processing may be important for the mastery of initial decoding skills but higher levels of fluent reading may depend more on simultaneous processing. In other words, among normal readers the conceptual-linguistic operations necessary for fluent reading may become independent of successive processing and consequently reading performance may relate more to simultaneous than to successive processing. However, among children who experience reading difficulties, initial deficits in successive processing may delay the differentiation of conceptual-linguistic operations from more elementary forms of sequential linguistic processing. If this is the case, we would expect that simultaneous and successive processing would relate differently to reading performance in low-achieving and normal readers. There are several studies which bear on this issue.

#### *Reading Performance and Simultaneous-Successive Processing*

Among RD and EMR children, reading difficulties have frequently been related to deficits in sequential processing. Doehring (1968) and Kinsbourne and Warrington (1966), for example, have reported that reading disability was significantly related to performance on tasks involving sequential processing. Blackman and his colleagues (Blackman, Bilsky, Burger & Mar, 1976; Blackman & Burger, 1972) have reported that sequential memory processes were highly related to the acquisition of word recognition skills among EMR children. Blackman et al., (1976) suggest that

those few mentally retarded subjects capable of deliberately generating strategies and applying them to the problem of recalling random digit strings may be able to generate and apply comparable strategies to the memory requirements — particularly those involved in decoding — of learning to read. (p. 132)

However, they suggest that although memory variables may be central to the acquisition of early reading skills in EMR populations, more complex cognitive variables may assume greater importance at more advanced levels of skills acquisition, such as comprehension in reading.

We have recently related simultaneous and successive processing to academic achievement in a group of 52 EMR children between the ages of 10 and 14. Simultaneous and successive factor scores derived from a factor analysis of four of the six marker tests in the Das battery (see Das, Kirby & Jarman, 1975) were correlated with the spelling, arithmetic, and oral reading subtests of the Wide Range Achievement Tests (WRAT) and with the Schonell silent reading subtest. These correlations are shown in Table 1.

It can be seen that successive processing correlates significantly with the WRAT oral reading and spelling subtests while simultaneous processing is significantly related to WRAT arithmetic. Silent reading which involves comprehension is related to neither processing strategy. These findings are



TABLE 1  
CORRELATIONS BETWEEN ACHIEVEMENT AND SIMULTANEOUS-  
SUCCESSIVE PROCESSING IN EMR CHILDREN

	Simultaneous	Successive
WRAT Spelling	-.21	.32*
WRAT Arithmetic	.28*	-.03
WRAT Oral Reading	-.17	.33*
Schonell Silent Reading	.09	.16

\*  $p < .05$

consistent with those of Blackman and his colleagues and suggest the importance of successive processing strategies for the development of decoding skills in EMR children.

Several other studies have related the Das battery of simultaneous and successive tasks to reading performance. Das, Manos and Kanungo (1975) compared the contributions of simultaneous and successive processing to reading achievement in high and low SES children. Briefly stated, the results indicated that the low SES group predominantly used successive processing as far as reading was concerned, whereas the high SES groups used a mixture of both processes. A recent study carried out by Hans Sprecht (unpublished) found that in a group of 65 low achieving high school students, reading achievement was significantly related to successive but not simultaneous tasks.

These findings suggest that low levels of successive processing ability may be a limiting factor in the reading achievement of RD, EMR, low SES and low achieving children. Krywaniuk (1974) and Kaufman (in preparation) have shown that teaching successive processing strategies to groups of poor readers led to significant improvements in decoding skills. It should not be inferred from these studies, however, that simultaneous processing is unrelated to reading performance. Satz (1976) has pointed out that "in early reading considerable weight is given to perceptual operations from which meaning is deduced from surface structure. By contrast, in later stages fluent reading is aided more by linguistic operations that sample the visual information to conform to prior expectations concerning meaning" (p. 279). Since, according to Luria, simultaneous processing is necessary for the discovery of conceptual relationships between objects and events, one would expect it to be involved at more advanced stages of reading. In studies involving children who are not experiencing reading difficulties we have indeed found this to be the case.

Kirby and Das (1977) divided a sample of 99 grade 4 children into four quadrants on the basis of median splits of simultaneous and successive factor scores and compared the performance of the four groups on the vocabulary and comprehension parts of the Gates McGinitie Reading Test. The results clearly showed that children high in both simultaneous and successive modes of processing had the highest level of reading achievement and children low in both modes had the lowest. The reading achievement of children high in one mode and low in the other was intermediate between the two extreme

groups and did not differ from each other. Kirby and Das conclude that high levels of simultaneous and successive processing are both necessary, though neither by itself is sufficient for high reading achievement.

We have recently followed up these findings with a study involving different criterion measures of reading achievement. Sixty grade 3 students were involved and the analytic procedure employed by Kirby and Das was again followed. Subjects were divided into four groups on the basis of simultaneous and successive factor scores and performance of the groups on the Decoding and Comprehension sections of the Elementary Reading Test developed by the Edmonton Public School Board was compared. The means and standard deviations of the four groups are shown in Table 2.

TABLE 2  
MEANS AND STANDARD DEVIATIONS OF FOUR GRADE 3 GROUPS  
ON READING SCORES

	Decoding	Comprehension
Hi Sim Hi Succ	50.27 (5.34) <sup>a</sup>	56.40 (12.12)
Hi Sim Lo Succ	52.13 (4.61)	56.44 (9.54)
Lo Sim Hi Succ	48.27 (4.70)	51.73 (10.12)
Lo Sim Lo Succ	47.13 (5.50)	49.07 (10.89)

<sup>a</sup> Standard deviations are given in parentheses.

Two-by-two analyses of variance showed significant main effects for simultaneous processing in both Decoding ( $F = 7.20$ ,  $df\ 1,56$ ,  $p < .01$ ) and Comprehension ( $F = 4.65$ ,  $df\ 1,56$ ,  $p < .05$ ) analyses. The main effect for successive processing and the interaction were nonsignificant in both analyses. The results of the analyses of variance were supported by correlations between simultaneous and successive factor scores and reading performance. Simultaneous factor scores correlated significantly with both Decoding ( $r = .45$ ,  $p < .001$ ) and Comprehension ( $r = .40$ ,  $p = .001$ ) performance. Correlations with successive factor scores were nonsignificant.

In order to further investigate the hypothesis that simultaneous processing is especially important at more advanced stages of reading we performed median splits on both comprehension and decoding distributions and carried out correlational analyses in the top and bottom halves of each distribution. Despite the reduced variance in comprehension scores, simultaneous processing correlated significantly ( $r = .41$ ,  $p < .05$ ) with comprehension in the top, but not the bottom, half of the distribution. When the decoding distribution was divided into top and bottom halves the relationship between decoding and simultaneous processing disappeared. From this it can be concluded that simultaneous processing may be necessary for the development of more advanced levels of comprehension skills. The fact that in the present study successive processing related less to reading performance than in the Kirby and Das study may be due either to the different criterion



measures of reading employed or to the fact that the present sample had a somewhat higher mean IQ than the Kirby and Das sample.

In summary, what emerges from these studies is that, among children who are likely to experience difficulty in reading, competence in successive processing is strongly related to reading achievement. However, among normal readers at more advanced levels of reading skills, simultaneous processing is equally, if not more, important in the reading process. These studies suggest the potential of the simultaneous/successive model as a framework for investigating the cognitive processing of children who experience reading difficulties. Understanding of the underlying processing deficits of these children is essential for effective and rationally-based remediation. The simultaneous/successive model is especially useful because of its explicit neuropsychological reference. Luria's observations have specified the cortical locations underlying different types of simultaneous and successive processes. Thus, differences between reading disabled children and normal readers in the relationships between simultaneous and successive dimensions and aspects of linguistic and reading performance can be explicitly related to underlying neurological phenomena. In order to illustrate how the model is relevant in investigating the problems of reading disabilities, we shall review some of the recent research on the cognitive characteristics of RD children.

### *Cognitive Characteristics of Reading Disabled Children*

Researchers are in substantial agreement that a unitary model of reading disability is unlikely to account for the observed phenomena. Rather, there are probably many different types of reading disability with many different causes. Doehring (1976) points out that reading acquisition itself is not a unitary process but rather involves a number of component skills. Consequently, different types of reading disability may be related to difficulties in acquiring different component skills. If the component reading skills are interdependent, however, deficiencies in one or two skills could lead to a general retardation in all reading skills.

Despite the complexity of the issue several broad groups of RD children can be distinguished. Three subgroups of RD children have been identified by several investigators. The first, and largest, subgroup consists of children who show a deficit, or possibly a lag (Satz, 1976), in auditory-temporal processing. Children in this subgroup generally have lower verbal than performance WISC IQ's and are often characterized by speech and language problems. The importance of language-related problems in reading disability can be seen in the fact that 83% of Warrington's (1967) sample of RD children performed more poorly on the verbal scale of the WISC than on the performance scale and a major subgroup of RD children identified by Mattis, French and Rapin (1975) was characterized by problems related to general language acquisition. There is evidence that the linguistic problems of this subgroup may stem from difficulties in successive processing related to neurolinguistic factors associated with the dominant fronto-temporal lobe. Doehring (1968), for example, reported that reading disability was most highly correlated with tasks that required sequential processing; similar relationships between sequential ordering difficulty and reading disability



were also found by Kinsbourne and Warrington (1966). Tallal (1976) has also reported deficiencies in auditory-temporal processing among RD children.

This subgroup of “auditory dyslexics” is contrasted by several investigators (e.g., Myklebust, 1965; Tallal, 1976) with a “visual dyslexic” subgroup comprised of children who have difficulty analyzing the visual-spatial aspects of written language. In Johnson and Myklebust’s (1964) study, 43% of RD children were reported to have visual processing disturbances. However, the proportion of “visual dyslexic” children was considerably less in Warrington’s (1967) study. Doehring (1968) and Ingram (1960) report that this group of RD children with visual processing difficulties were characterized by performance IQ deficits on the WISC, reversals in letters and words and by directional confusion and poor lateralization. In terms of Luria’s model these children are clearly experiencing difficulties in simultaneous processing.

Boder (1973) has suggested that in addition to RD children with auditory-temporal and visual-spatial difficulties there exists a third subgroup which has a more pervasive deficit incorporating difficulties in both auditory and visual processing. Doehring (1968) also distinguished a third group of RD children which was characterized by deficits in both auditory and visual modalities. In addition to *specific* processing deficits, many RD children are characterized by arousal problems (see Das, Leong & Williams, 1977). However, here we are concerned only with specific processing deficits.

Clearly the nature of learning and reading disabilities is multifaceted and studies conducted within the simultaneous-successive framework support the multifaceted nature of these disabilities. Williams (1977), for example, found significant differences between learning disabled and IQ-matched normal children on successive processing tasks and on Figure Copying — a measure of simultaneous processing. Comparison of these learning disabled children with a different control group replicated these findings (Das et al., 1977). Leong (1976) found highly significant differences between RD and normal children on all the simultaneous and successive tasks, despite the fact that groups were matched for IQ. In both these studies factor analyses of simultaneous and successive tasks revealed similar factor patterns in RD and control groups. Das et al. (1977) conclude that although learning disabled and normal children use similar processes in solving the tasks of the battery, the learning disabled children are considerably less efficient in the use of both simultaneous and successive processes.

Satz and his associates (Satz & Sparrow, 1970; Satz & Van Nostrand, 1973) have suggested that developmental dyslexia is a result of a lag in the maturation of the left hemisphere which differentially delays the development of perceptual and conceptual-linguistic abilities as a function of their ontogenetic sequence of development. This theory assumes that perceptual operations are crucial in developing initial reading skills but conceptual-linguistic operations are considerably more important at later stages. The high incidence of WISC Verbal IQ deficiencies associated with reading disability testifies to the importance of conceptual-linguistic factors. Studies of dichotic listening (e.g., Leong, 1976) also provide partial support for the hypothesis of a lag in the development of left hemisphere among reading disabled children. However, as Satz (1976) points out, the finding of a lag in the magnitude of ear asymmetry in disabled readers “has been neither robust nor reliable.”

Rourke (1975, 1976) has disputed the "developmental lag" hypothesis, arguing instead that the cognitive difficulties of RD children constitute a deficit rather than a lag. He points out that the deficit view "differs from the 'developmental lag' position in that there is no necessary expectation that the children who suffer from the deficit(s) will ever catch up with their normal age-mates in those skills required for age-appropriate reading" (1976, pp. 126-127). In support of the deficit position, Rourke (1975) cites the studies of Reitan and his associates (e.g., Reed & Reitan, 1963; Reitan, 1955) who demonstrated a relationship between left hemisphere lesions and selective impairment on the verbal scales of the Wechsler-Bellevue on the one hand, and right hemisphere lesions and selective impairment on the performance scales on the other. The implication is that verbal skills are mediated principally by the left hemisphere and visual-spatial skills by the right hemisphere. Rourke has shown that older learning disabled children (9- to 14-year-old) who exhibit specific patterns of Verbal IQ-Performance IQ discrepancies on the WISC "do behave in a manner that is quite similar to that which would be expected (on the basis of adult data) if they were experiencing the effects of cerebral dysfunction" (1975, p. 916). However, younger learning disabled children (5- to 8-year-old) do not exhibit the same clear cut patterns of abilities and deficits. Thus, the developmental aspects of learning disabilities are crucial (Rourke, 1976).

Thus far, our review of the cognitive characteristics of RD children has indicated the multifaceted nature of the phenomenon. Reading difficulties can result from a variety of processing and arousal problems. However, it is generally agreed that a large proportion of reading difficulties involve problems in general language acquisition possibly deriving from a deficit in successive processing. Several investigators (e.g., Rourke, 1976; Satz, 1976) have emphasized the importance of a developmental perspective since processing requirements of later fluent reading may differ from those of initial reading acquisition. Satz (1976), for example, points out that at more advanced levels, fluent reading performance depends not only on word recognition skills but also on the ability to process and anticipate semantic information. Thus, a relevant question for understanding the nature of reading disability concerns the developmental interrelationships between reading performance, sequencing abilities and semantic or conceptual-linguistic processing.

### *Reading Disability and the Development of Conceptual-Linguistic Processing*

Patel (1976, in press) has recently argued for the centrality of conceptual-linguistic difficulties for some subgroups of reading disabled children. Following Mattis, French and Rapin (1975) Patel points out that one major type of reading disability involves problems related to general language acquisition. He suggests that reading disabled children may have difficulty in processing the relational aspects of words or, in Luria's terms, logico-grammatical relations. In support of this hypothesis he cites Vogel's (1975) finding of a large discrepancy between dyslexic and normal children in the extent to which semantic factors accounted for unique variance in reading comprehension scores. Vogel reported, as did Doebring (1968), that receptive vocabulary correlated significantly with reading comprehension only in



normal readers. In RD children the correlation was close to zero. Vogel points out that this discrepancy is all the more striking in view of the lack of any significant differences between RD and normal readers in receptive vocabulary. She suggests that the RD child "because of his syntactic deficiencies is unable to comprehend the relational aspect of words and therefore is blocked from fully utilizing the semantic information he possesses" (p. 76).

Patel (1976, in press) points out that the semantic-syntactic systems which develop in children around age seven (e.g., Chomsky, 1969; Goodglass, Berko & Hyde, 1970; Kessel, 1970) are similar to the linguistic disorders observed in semantic aphasia by Luria (1970, 1976). Luria has shown that semantic aphasics who have suffered lesions to the parieto-temporo-occipital junction are unable to process the relational aspects of words. They cannot relate the meanings of the words they understand in isolation. For example, anomalous phrases such as "the wife of the sister" would not be noticed by these patients. Processing of spatio-temporal prepositions (e.g., over, before, in front of, in back of, etc.) and passive sentences are also impaired in semantic aphasics. According to Luria, performance of these tasks requires simultaneous processing of logical relations and depends on integrity of the parieto-temporo-occipital junction, which develops around age seven (Luria, 1970).

On the basis of these observations, Patel (1976) suggests that some reading disabled children may not have "acquired the ability to process the relational aspect of words (logico-grammatical relations) due to some difference in their neurolinguistic mechanisms associated with the parieto-temporo-occipital junction" (p. 305). Patel's hypothesis is consistent with the high incidence of verbal IQ deficits among RD children. However, because it is based on Luria's observations, the hypothesis is extremely powerful in terms of specifying brain-behaviour relationships. The hypothesis also raises important research questions regarding the development of reading disabilities which can be posed within the framework of the simultaneous-successive processing model.

For example, to what extent are the hypothesized conceptual-linguistic or logical-grammatical deficits in RD children just one aspect of a more pervasive sequencing deficit, or alternatively, are there two distinct subgroups of RD children (within the more general "language difficulties" subgroup), one suffering from sequencing difficulties and the other from conceptual-linguistic difficulties? Based on Luria's neuropsychological observations, it would be predicted that performance on logical grammatical tasks in normal adults or adolescents would be more closely related to simultaneous than to successive processing. However, at younger age levels when these semantic systems are just beginning to emerge, this relationship may not hold. In fact, the prerequisite skills for the processing of linguistic input are likely to be more dependent on successive than on simultaneous processing. In children who have difficulties in successive processing, this may be especially so. The studies relating simultaneous and successive processing to reading performance in various groups support this hypothesis.

Thus RD children who are characterized by a deficit in successive processing may experience a lag in the differentiation of deeper levels of conceptual-linguistic abilities from more elementary or surface forms of



sequential-linguistic processing. Since fluent reading skills depend heavily on semantically-based linguistic operations this delayed differentiation may explain the significant relationships which have been observed between successive processing and reading performance among RD and other low-achieving groups. However, among normal readers, conceptual-linguistic operations may have become more independent of successive processing with the result that simultaneous processes play a greater role in explaining variance in reading performance.

In this paper our aim has been to show the relevance of the simultaneous-/successive model of cognitive processing for research in reading disabilities. Studies relating simultaneous and successive processing to reading suggest that, among less fluent readers, successive processing accounts for more variance in reading scores than does simultaneous processing. However, among more fluent readers, simultaneous processes assume as much, if not more, importance as successive processes. It is generally agreed that one major form of reading disability involves problems relating to general language acquisition, probably stemming from deficiencies in successive or auditory-temporal processing. It has also been argued that RD children have difficulty in processing the semantic information which is especially important in later fluent reading. Within the framework of the simultaneous-/successive model, we have raised a variety of research questions regarding the developmental interrelationships between reading performance, conceptual-linguistic skills and cognitive processing in both clinical and normal groups of children. We believe that it is only on the basis of this type of research into the underlying processing deficits of disabled readers that rationally-based remediation procedures can be implemented.

Preparation of this paper was made possible by a grant to the Centre for the Study of Mental Retardation by the Edmonton Civic Employees Welfare Chest Fund.

### References

- Blackman, L. S., & Burger, A. L. Psychological factors related to early reading behaviour of EMR and nonretarded children. *American Journal of Mental Deficiency*, 1972, 77, 212-229.
- Blackman, L. S., Bilsky, L. H., Burger, A. L., & Mar, H. Cognitive processes and academic achievement in EMR adolescents. *American Journal of Mental Deficiency*, 1976, 81, 125-134.
- Boder, E. Developmental dyslexia: A diagnostic approach based on three reading-spelling patterns. *Developmental Medicine and Child Neurology*, 1973, 15, 663-687.
- Chomsky, C. *The acquisition of syntax in children from 5 to 10*. Cambridge, Mass.: M.I.T. Press, 1969.
- Cummins, J. Systems of mediation in memory and reasoning. Paper presented to the annual conference of the Canadian Psychological Association, Victoria, June, 1973.
- Cummins, J., & Das, J. P. Simultaneous and successive syntheses and linguistic processing. Manuscript submitted for publication, 1977.
- Das, J. P., Kirby, J., & Jarman, R. F. Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin*, 1975, 82, 87-103.
- Das, J. P., Leong, C. K., & Williams, N. H. Relation between learning disability and simultaneous-successive processing. Paper submitted for publication, 1977.

- Das, J. P., Manos, J., & Kanungo, R. N. Performance of Canadian native, black and white children on some cognitive and personality tests. *Alberta Journal of Educational Research*, 1975, 21, 183-195.
- Doehring, D. G. *Patterns of impairment in specific reading disability*. Bloomington, Ind.: Indiana University Press, 1968.
- Doehring, D. G. Evaluation of two models of reading disability. In R. M. Knights & D. J. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches*. Baltimore: University Park Press, 1976.
- Goodglass, H., Berko, J., & Hyde, M. Some dimensions of auditory language comprehension in aphasia. *Journal of Speech and Hearing Research*, 1970, 13, 595-606.
- Ingram, T. T. S. Pediatric aspects of specific developmental dysphasia, dyslexia and dysgraphia. *Cerebral Palsy Bulletin*, 1960, 2, 254-267.
- Johnson, D., & Myklebust, H. *Learning disabilities: Educational principles and practices*. New York: Grune and Stratton, 1964.
- Kaufman, D. The influence of learning strategies on academic performance (in preparation).
- Kessel, F. The role of syntax in children's comprehension from age six to twelve. *Monographs of the Society for Research in Child Development*, 1970, 139.
- Kinsbourne, M., & Warrington, E. K. Developmental factors in reading and writing backwardness. In J. Money (Ed.), *The disabled reader*. Baltimore: The Johns Hopkins Press, 1966.
- Kirby, J. R., & Das, J. P. Reading achievement, IQ, and simultaneous-successive processing. *Journal of Educational Psychology*, 1977, 69, 564-570.
- Krywaniuk, L. W. *Patterns of cognitive abilities of high and low achieving school children*. Unpublished doctoral dissertation, University of Alberta, 1974.
- Leong, C. K. Lateralization in severely disabled readers in relation to functional cerebral development and synthesis of information. In R. M. Knights & D. J. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches*. Baltimore: University Park Press, 1976.
- Luria, A. R. *Higher cortical functions in man*. New York: Basic Books, 1966.
- Luria, A. R. *Traumatic aphasia: Its syndromes, psychology and treatment*. The Hague: Mouton, 1970.
- Luria, A. R. *The working brain: An introduction to neuropsychology*. Harmondsworth: Penguin, 1973.
- Luria, A. R. Two kinds of disorders in the comprehension of grammatical constructions. *Linguistics*, 1976, 47-56.
- Mattis, S., French, J. H., & Rapin, I. Dyslexia in children and young adults: Three independent neuropsychological syndromes. *Developmental Medicine and Child Neurology*, 1975, 17, 150-163.
- Myklebust, H. R. *Development and disorder of written language: Picture story language test*. New York: Grune & Stratton, 1965.
- Patel, P. G. Review of "Syntactic abilities in normal and dyslexic children." Baltimore: University Park Press, 1975. *Journal of Child Language*, 1976, 3, 303-307.
- Patel, P. G. The left parieto-temporo-occipital junction, semantic aphasia and language development around age seven. *Linguistics*, (in press).
- Reed, H. B. C., & Reitan, R. M. Intelligence test performance of brain-damaged subjects with lateralized motor deficits. *Journal of Consulting Psychology*, 1963, 27, 102-106.
- Reitan, R. M. Certain differential effects of left and right cerebral lesions in human adults. *Journal of Comparative and Physiological Psychology*, 1955, 48, 474-477.
- Rourke, B. P. Brain-behaviour relationships in children with learning disabilities. A research program. *American Psychologist*, 1975, 30, 911-920.
- Rourke, B. P. *Reading retardation in children: Developmental lag or deficit?* In R. M. Knights & D. J. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches*. Baltimore: University Park Press, 1976.

- Satz, P. Cerebral dominance and reading disability: An old problem revisited. In R. M. Knights & D. J. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches*. Baltimore: University Park Press, 1976.
- Satz, P., & Sparrow, S. Specific developmental dyslexia: A theoretical formulation. In D. J. Bakker & P. Satz (Eds.), *Specific reading disability: Advances in theory and method*. Rotterdam: Rotterdam University Press, 1970.
- Satz, P., & Van Nostrand, G. K. Developmental dyslexia: An evaluation of a theory. In P. Satz & J. J. Ross (Eds.), *The disabled learner: Early detection and intervention*. Rotterdam: Rotterdam University Press, 1973.
- Tallal, P. Auditory perceptual factors in language and learning disabilities. In R. M. Knights & D. J. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches*. Baltimore: University Park Press, 1976.
- Vogel, S. A. *Syntactic abilities in normal and dyslexic children*. Baltimore: University Park Press, 1975.
- Warrington, E. K. The incidence of verbal disability associated with retardation reading. *Neuropsychologia*, 1967, 5, 175-179.
- Williams, N. H. Learning disability. Paper presented to the annual conference of the Canadian Psychological Association, Vancouver, 1977.



LAYMAN E. ALLEN

and

JOAN ROSS

*University of Michigan*

## Improving Skill in Applying Mathematical Ideas: A Preliminary Report on the Instructional Gaming Program at Pelham Middle School in Detroit

*Performance of 237 students in ten inner-city eighth-grade math classes was assessed under five different sets of instructional conditions to measure their ability to solve two types of problems: (1) those that involve only computation and (2) those that involve not only computation, but also recognition of the relevance of a particular mathematical idea that is indispensable to the solution of the problem. Results indicate that the combination of playing EQUATIONS (an instructional mathematics game) over a two-year period and then working intensively with IMP (Instructional Math Play) Kits for two weeks enables students to apply mathematical ideas (in the sense studied in this experiment) better (at the .0001 level of significance) than any of the other four sets of conditions: (1) playing EQUATIONS alone, (2) playing EQUATIONS and being taught explicitly by a teacher the ideas presented in the IMP Kits, (3) being taught the ideas in an ordinary math class without playing EQUATIONS, and (4) participating in an ordinary math class without explicit teaching of the ideas or playing EQUATIONS. (Mr. Allen is Professor of Law and Research Scientist at the Law School and the Mental Health Research Institute at the University of Michigan; Mrs. Ross is Research Associate at the Mental Health Research Institute.)*

Educators working with the EQUATIONS game and associated materials quickly become aware that applying mathematical ideas is much more difficult than merely computing with those ideas. Of course, "applying" an idea is a somewhat vague notion; this preliminary report will be addressed to one aspect of application — namely, recognition that an idea is indispensably relevant to the solution of a problem. Consider the following pair of problems: the first is the C-type (computation), and the second is the R-type (relevance).

1.  $6 - (1-3) = \underline{\hspace{1cm}} ? \underline{\hspace{1cm}}$ .

	<u>A</u>	<u>B</u>	<u>C</u>	<u>Yes</u>	<u>No</u>
2.	- - 1 3 6	8	<u>          </u>	<u>      </u>	<u>      </u>

By writing an X in the Yes or No column, indicate whether or not all of the numbers and operations in Column A can be appropriately ordered and grouped (inserting parentheses wherever necessary) to form an expression equal to the number in Column B. If your answer is Yes, write that expression in Column C.

The indispensably relevant idea for solving each of the problems is the subtraction of negative numbers. In the C-type problem, the very statement of the problem clearly and explicitly indicates that subtracting a negative number is involved. That is neither so clearly nor so explicitly cued in the statement of the R-type problem. Those who understand how to subtract negative numbers can easily do the first problem correctly. But many of those who can solve a C-type problem involving subtraction of negative numbers fail to solve a corresponding R-type problem involving negative numbers. In general (in the groups we have studied), about two-thirds of those who solve C-type problems fail to solve a corresponding R-type problem that involves the same idea. The R-type problem is considerably harder than the C-type in this example because the student must recognize from less clear and less explicit cues that subtracting negative numbers is an indispensably relevant idea for solving the problem. “Understanding” an idea in the R-sense (being able to solve R-type problems) includes understanding it in the C-sense, but it also involves something more. R-sense understanding includes the capability of selecting from among a storehouse of ideas understood in the C-sense, those that are indispensably relevant for solving a particular problem. The question to which this study is addressed is whether skills in applying mathematical ideas can be improved by learning procedures which emphasize exposure to situations that are rich in opportunities for such application, at levels of complexity appropriate for each learner.

*The EQUATIONS Game and the IMP (Instructional Math Play) Kits*

The rules that define the EQUATIONS game establish a problem-generating and problem-solving interaction between small groups of students, an interaction that can easily be controlled to provide a highly individualized learning experience for each of the participants. It is a RAG (Resource Allocation Game) where the resources involved are mathematical ideas. (For details, see Allen, 1972.) The IMP Kits are 16-page pamphlet-simulations of a computer playing EQUATIONS where the computer is programmed to play like a good teacher, rather than like a good player. Each kit presents a lesson on one mathematical idea. For example, in the following situation.

RESOURCES:	+ - - 1 3 3 ( )	FORBIDDEN: 2
	=	PERMITTED:
<u>SOLUTION</u>	<u>5</u>	REQUIRED:
	GOAL	

on its turn to move, the computer might move the + from RESOURCES to FORBIDDEN, thus extinguishing the SOLUTION  $(3 + 3) - 1$  and presenting the learner, in effect, with the question:

Is it still possible (after the + is FORBIDDEN) to construct an expression equal to the GOAL of 5 from the remaining RESOURCES?

If the learner by challenging the computer's move, in effect, answers "NO," then the computer will direct the learner to a comment that teaches a lesson in the subtraction of negative numbers:

Your challenge that all SOLUTIONS have been extinguished is incorrect. The SOLUTION  $3 - (1-3)$  is still possible. Notice how a pair of minus signs can be used to give the effect of addition. Since  $3-5 = -2$  and  $3 - -2 = 5$ , the GOAL can be achieved even though the + is FORBIDDEN. Go on to the next IMP Kit.

There are at present five versions of each of 21 ideas, or a total of 105 kits. The first 21 of these IMP Kits have been published and are available to those interested. (For details, see Allen & Ross, 1975.)

### *Method*

#### *Subjects*

All of the students in each of ten of the fourteen eighth-grade classes in mathematics at Pelham Middle School participated in the study. The ten classes were chosen to include all four of the classes in which the EQUATIONS game had been used during the prior two years as part of the regular instructional program in mathematics and two other classes of each of the three participating teachers. Pretest and/or posttest data were collected on 237 of the students enrolled in these ten classes.

#### *Experimental Treatments*

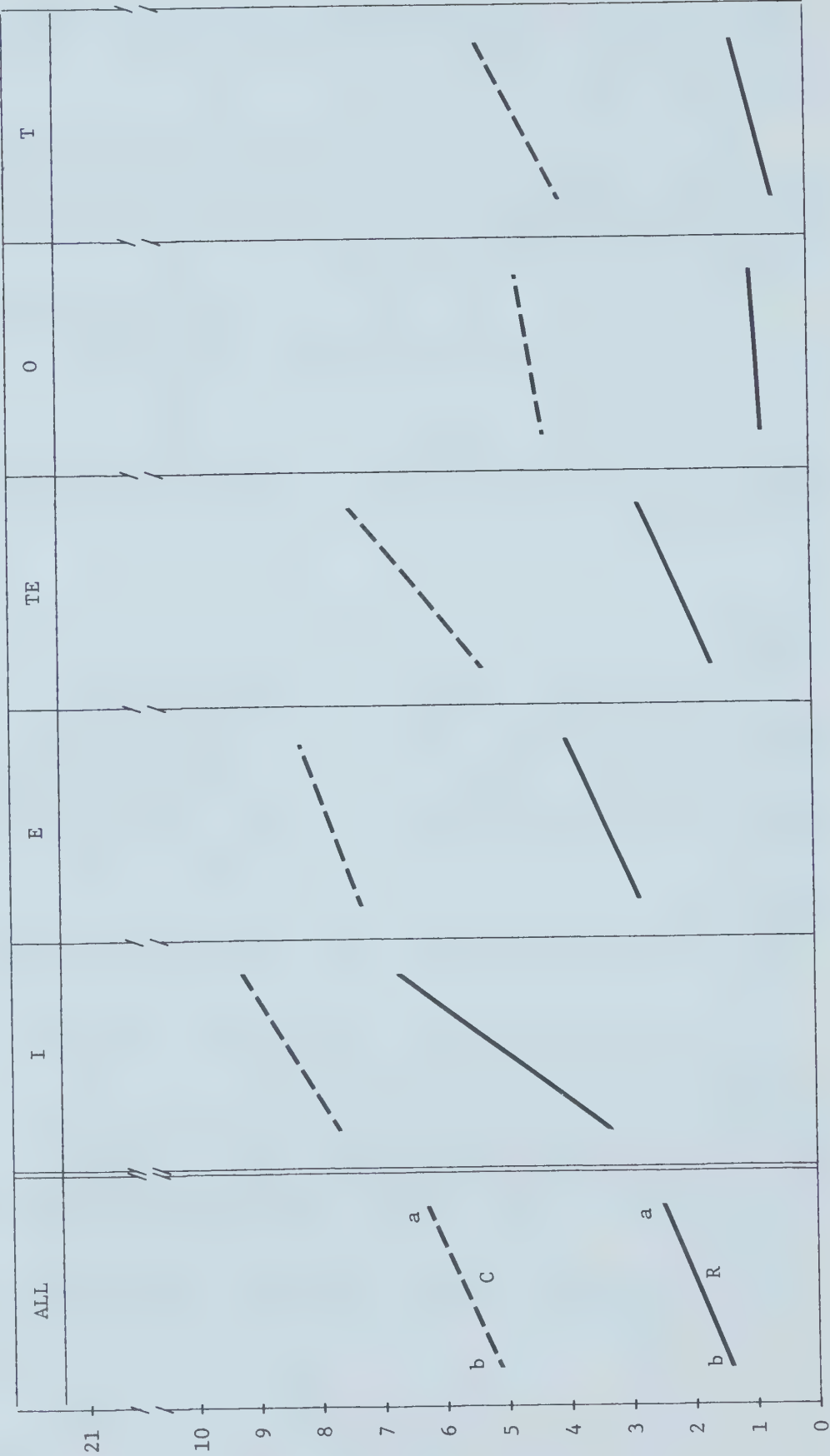
The following five different sets of experimental conditions were represented in the ten classes:

- I an EQUATIONS class in which the IMP Kits were used in five class periods during the two-week experiment and the regular once-a-week EQUATIONS tournament was continued;
- E an EQUATIONS class in which the game was played for the five class periods without any explicit teaching of the 21 IMP Kit ideas;
- TE two EQUATIONS classes in which the game was played for five class periods and the teachers explicitly taught the 21 ideas presented in the IMP Kits;
- O three non-EQUATIONS classes in which the ordinary classroom procedure was continued without change with no special attention given to the 21 IMP Kit ideas; and
- T three non-EQUATIONS classes in which the teachers explicitly taught the 21 IMP Kit ideas for five class periods.

The set of conditions of greatest interest to the researchers was that of the I group in which students individually played through the IMP Kits, completing as many of the set of 105 as they could in the five periods. Before



Figure 1. Mathematics pretest and posttest scores on C Tests and R Tests of five experimental groups of eighth-grade classes, Pelham Middle School, Detroit, 1974.



	b	a	a - b	b	a	a - b	b	a	a - b	b	a	a - b	b	a	a - b
C	206	202	188	23	23	23	18	17	17	46	44	44	58	61	54
N	5.10	6.32	1.26	7.70	9.26	1.57	7.39	8.35	0.82	5.33	7.48	2.07	4.31	4.74	0.44
$\bar{X}$	2.50	3.23	2.36	2.29	2.75	2.39	2.70	2.87	2.40	2.33	3.69	2.43	1.81	2.20	2.28
Sx			.0001			.005			ns			.0001			ns
Signif.															
R	207	202	190	23	23	23	17	17	16	46	44	44	60	60	56
N	1.40	2.43	1.03	3.39	6.78	3.40	2.88	4.12	1.00	1.65	2.84	1.11	0.83	0.95	0.11
$\bar{X}$	1.70	2.79	1.81	1.56	3.16	2.27	2.00	2.32	1.93	1.83	2.62	1.79	1.15	1.21	0.95
Sx			.0001			.0001			ns			.0005			ns
Signif.															
C - R	202	198	181	23	23	23	17	17	16	46	44	44	58	59	53
N	3.73	3.93	0.20	4.30	2.48	-1.83	4.59	4.24	-0.19	3.67	4.64	0.95	3.45	3.85	0.40
$\bar{X}$	1.88	2.22	2.60	1.99	1.86	2.35	2.83	2.17	3.74	1.71	2.51	2.56	1.62	1.90	2.12
Sx			ns			.005			ns			.05			ns
Signif.															
ALL					I			E			TE			O	
															T

KEY

Tests:

- C test  
----- R test  
a (after) posttest  
b (before) pretest  
C C test  
R R test

Treatments:

- I IMP Kits  
E EQUATIONS  
TE Taught 21 Ideas, EQUATIONS  
O Ordinary classroom activities  
T Taught 21 Ideas

this group started on the IMP Kits, one class period was devoted to teaching members of the class how to use the kits.

*Dependent Variables*

The effects of the various sets of experimental conditions were measured by two different forms of a pair of specially-constructed tests targeted at the 21 mathematical ideas presented in the IMP Kits. The first of the pair of tests is called a C test; it contains only C-type items. The second test is called an R test; it contains only R-type items. Two different forms of the C test were used (Form C and Form D), as well as two different forms of the R test (Form E and Form F). In each of the ten classes in which these tests were administered, the students were divided into eight groups — G1, G2, . . . G8. Each student received a C test and R test as pretests, and each received alternative forms of the two tests as posttests as follows:

	<i>Order of Administration</i>	<i>Groups</i>							
		<i>G1</i>	<i>G2</i>	<i>G3</i>	<i>G4</i>	<i>G5</i>	<i>G6</i>	<i>G7</i>	<i>G8</i>
Pretest	1	C	C	E	F	D	D	E	F
	2	E	F	C	C	E	F	D	D
Posttest	1	F	E	D	D	F	E	C	C
	2	D	D	F	E	C	C	F	E

Using Ca (after) to denote the score on the C posttest and Cb (before) to denote the C pretest score (and similarly for the R pretests and posttests), outcome measures of three dependent variables can be specified as follows:

- |                      |  |
|----------------------|--|
| 1. Ca-Cb             | Increase in performance on C test                          |
| 2. Ra-Rb             | Increase in performance on R test                          |
| 3. (Ca-Ra) - (Cb-Rb) | Decrease in difference in performance on C test and R test |

*Results*

The scores for each of the five experimental groups, summarized in Figure 1, were significantly higher (at the .0001 level) on the C test than on the R test both on pretests and on posttests. The mean pretest score for all students on the C test was 5.10, while for the R test it was 1.40 (maximum score = 21), a ratio of about 3.6 to 1. On the posttests the ratio decreased to 2.6 to 1 with mean scores of 6.32 and 2.43, respectively.

Three of the experimental groups had significant differences between pretest and posttest scores on the C test (measured by Ca-Cb). The IMP Kit group (I) had a mean pretest of 7.70 and 9.26 on the posttest, significantly higher at the .001 level, and the groups that were explicitly taught the 21 ideas (TE and T) went from 5.33 to 7.48 and from 4.02 to 5.32, respectively, significantly higher on the posttest at the .0001 level.

The same three experimental groups had significant differences between pretest and posttest scores on the R test (measured by Ra-Rb). The IMP Kit group and the T group were significantly higher on the posttest (.0001) with mean pretest to posttest scores of 3.39 to 6.78 and 0.59 to 1.41, respectively, whereas the TE group was significantly higher at the .0005 level with scores of 1.65 to 2.84.



Only two of the experimental groups showed significant changes in the gap between C-sense understanding and R-sense understanding from the pretests to the posttests [as measured by (Ca-Ra) - (Cb-Rb)]. The IMP Kit group achieved a 1.83 reduction in its CR Gap, significant at the .005 level, in moving from a pretest gap of 4.30 to a posttest gap of 2.48. On the other hand, the T group increased its CR Gap by 0.95, significant at the .05 level, with a pretest gap of 3.67 and a posttest gap of 4.64.

In comparing the test scores of the experimental groups with each other, only those pairs in the total collection that qualify by the highly conservative Scheffé procedure (see Winer, 1971) at the .05 level of significance are reported as being significantly different. The results of the between-group comparisons are summarized in Figure 2.

Tests	Pretest	Posttest	Pretest - Posttest																																																																																										
C	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>T</th><th>O</th><th>TE</th><th>E</th><th>I</th></tr><tr><td>T</td><td>ns</td><td>.0005</td><td>.0001</td><td>.0001</td></tr><tr><td>O</td><td></td><td>ns</td><td>.0001</td><td>.0001</td></tr><tr><td>TE</td><td></td><td></td><td>.001</td><td>.0001</td></tr><tr><td>E</td><td></td><td></td><td></td><td>ns</td></tr></table>	Significance Lines and Levels					T	O	TE	E	I	T	ns	.0005	.0001	.0001	O		ns	.0001	.0001	TE			.001	.0001	E				ns	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>O</th><th>T</th><th>TE</th><th>E</th><th>I</th></tr><tr><td>O</td><td>ns</td><td>.0001</td><td>.0001</td><td>.0001</td></tr><tr><td>T</td><td></td><td>.0005</td><td>.0001</td><td>.0001</td></tr><tr><td>TE</td><td></td><td></td><td>ns</td><td>ns</td></tr><tr><td>E</td><td></td><td></td><td></td><td>ns</td></tr></table>	Significance Lines and Levels					O	T	TE	E	I	O	ns	.0001	.0001	.0001	T		.0005	.0001	.0001	TE			ns	ns	E				ns	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>O</th><th>E</th><th>T</th><th>I</th><th>TE</th></tr><tr><td>O</td><td>ns</td><td>ns</td><td>ns</td><td>.005</td></tr><tr><td>E</td><td></td><td>ns</td><td>ns</td><td>ns</td></tr><tr><td>T</td><td></td><td></td><td>ns</td><td>ns</td></tr><tr><td>I</td><td></td><td></td><td></td><td>ns</td></tr></table>	Significance Lines and Levels					O	E	T	I	TE	O	ns	ns	ns	.005	E		ns	ns	ns	T			ns	ns	I				ns
	Significance Lines and Levels																																																																																												
	T	O	TE	E	I																																																																																								
	T	ns	.0005	.0001	.0001																																																																																								
	O		ns	.0001	.0001																																																																																								
TE			.001	.0001																																																																																									
E				ns																																																																																									
Significance Lines and Levels																																																																																													
O	T	TE	E	I																																																																																									
O	ns	.0001	.0001	.0001																																																																																									
T		.0005	.0001	.0001																																																																																									
TE			ns	ns																																																																																									
E				ns																																																																																									
Significance Lines and Levels																																																																																													
O	E	T	I	TE																																																																																									
O	ns	ns	ns	.005																																																																																									
E		ns	ns	ns																																																																																									
T			ns	ns																																																																																									
I				ns																																																																																									
R	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>T</th><th>O</th><th>TE</th><th>E</th><th>I</th></tr><tr><td>T</td><td>ns</td><td>.0005</td><td>.0001</td><td>.0001</td></tr><tr><td>O</td><td></td><td>ns</td><td>.0001</td><td>.0001</td></tr><tr><td>TE</td><td></td><td></td><td>ns</td><td>.0001</td></tr><tr><td>E</td><td></td><td></td><td></td><td>ns</td></tr></table>	Significance Lines and Levels					T	O	TE	E	I	T	ns	.0005	.0001	.0001	O		ns	.0001	.0001	TE			ns	.0001	E				ns	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>O</th><th>T</th><th>TE</th><th>E</th><th>I</th></tr><tr><td>O</td><td>ns</td><td>.0001</td><td>.0001</td><td>.0001</td></tr><tr><td>T</td><td></td><td>.001</td><td>.0001</td><td>.0001</td></tr><tr><td>TE</td><td></td><td></td><td>ns</td><td>.0001</td></tr><tr><td>E</td><td></td><td></td><td></td><td>.0001</td></tr></table>	Significance Lines and Levels					O	T	TE	E	I	O	ns	.0001	.0001	.0001	T		.001	.0001	.0001	TE			ns	.0001	E				.0001	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>O</th><th>T</th><th>E</th><th>TE</th><th>I</th></tr><tr><td>O</td><td>ns</td><td>ns</td><td>.001</td><td>.0001</td></tr><tr><td>T</td><td></td><td>ns</td><td>ns</td><td>.0001</td></tr><tr><td>E</td><td></td><td></td><td>ns</td><td>.0001</td></tr><tr><td>TE</td><td></td><td></td><td></td><td>.0001</td></tr></table>	Significance Lines and Levels					O	T	E	TE	I	O	ns	ns	.001	.0001	T		ns	ns	.0001	E			ns	.0001	TE				.0001
	Significance Lines and Levels																																																																																												
	T	O	TE	E	I																																																																																								
	T	ns	.0005	.0001	.0001																																																																																								
	O		ns	.0001	.0001																																																																																								
TE			ns	.0001																																																																																									
E				ns																																																																																									
Significance Lines and Levels																																																																																													
O	T	TE	E	I																																																																																									
O	ns	.0001	.0001	.0001																																																																																									
T		.001	.0001	.0001																																																																																									
TE			ns	.0001																																																																																									
E				.0001																																																																																									
Significance Lines and Levels																																																																																													
O	T	E	TE	I																																																																																									
O	ns	ns	.001	.0001																																																																																									
T		ns	ns	.0001																																																																																									
E			ns	.0001																																																																																									
TE				.0001																																																																																									
C - R	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>E</th><th>I</th><th>TE</th><th>T</th><th>O</th></tr><tr><td>E</td><td>ns</td><td>ns</td><td>ns</td><td>ns</td></tr><tr><td>I</td><td></td><td>ns</td><td>ns</td><td>ns</td></tr><tr><td>TE</td><td></td><td></td><td>ns</td><td>ns</td></tr><tr><td>T</td><td></td><td></td><td></td><td>ns</td></tr></table>	Significance Lines and Levels					E	I	TE	T	O	E	ns	ns	ns	ns	I		ns	ns	ns	TE			ns	ns	T				ns	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>TE</th><th>E</th><th>T</th><th>O</th><th>I</th></tr><tr><td>TE</td><td>ns</td><td>ns</td><td>ns</td><td>.0001</td></tr><tr><td>E</td><td></td><td>ns</td><td>ns</td><td>ns</td></tr><tr><td>T</td><td></td><td></td><td>ns</td><td>ns</td></tr><tr><td>O</td><td></td><td></td><td></td><td>ns</td></tr></table>	Significance Lines and Levels					TE	E	T	O	I	TE	ns	ns	ns	.0001	E		ns	ns	ns	T			ns	ns	O				ns	<table><tr><th colspan="5">Significance Lines and Levels</th></tr><tr><th>TE</th><th>T</th><th>O</th><th>E</th><th>I</th></tr><tr><td>TE</td><td>ns</td><td>ns</td><td>ns</td><td>.0001</td></tr><tr><td>T</td><td></td><td>ns</td><td>ns</td><td>.0005</td></tr><tr><td>O</td><td></td><td></td><td>ns</td><td>.0005</td></tr><tr><td>E</td><td></td><td></td><td></td><td>ns</td></tr></table>	Significance Lines and Levels					TE	T	O	E	I	TE	ns	ns	ns	.0001	T		ns	ns	.0005	O			ns	.0005	E				ns
	Significance Lines and Levels																																																																																												
	E	I	TE	T	O																																																																																								
	E	ns	ns	ns	ns																																																																																								
	I		ns	ns	ns																																																																																								
TE			ns	ns																																																																																									
T				ns																																																																																									
Significance Lines and Levels																																																																																													
TE	E	T	O	I																																																																																									
TE	ns	ns	ns	.0001																																																																																									
E		ns	ns	ns																																																																																									
T			ns	ns																																																																																									
O				ns																																																																																									
Significance Lines and Levels																																																																																													
TE	T	O	E	I																																																																																									
TE	ns	ns	ns	.0001																																																																																									
T		ns	ns	.0005																																																																																									
O			ns	.0005																																																																																									
E				ns																																																																																									
<div>KEY</div> <div>Treatments:</div> <div>I IMP Kits</div> <div>E EQUATIONS</div> <div>TE Taught 21 Ideas, EQUATIONS</div> <div>O Ordinary classroom activities</div> <div>T Taught 21 Ideas</div>																																																																																													

Figure 2. Significance levels of pairs of sets of conditions that significantly differ from each other on test scores. (The only pairs of sets of conditions shown in this table as significantly different from each other are those in the collection of pairs that satisfy the highly conservative Scheffé procedure at the .05 level.)

On the pretest scores seven of the pairs of groups were significantly different on the C tests, six pairs were different on the R tests, but none were significantly different on the amount of the CR Gap. The significance lines written over the names of each of the experimental groups in Figure 2 should be interpreted as follows:

- 1. Groups whose names *do not* appear under a common line *do* differ significantly from each other;
- 2. Those whose names *do* appear under a common line *do not* differ significantly from each other.

Hence, on C pretest scores the IMP Kit group was significantly higher than the TE, O, and T groups (at the .0001 level); the E group was higher (.001) than the TE group and higher (.0001) than the O and T groups; and the TE group was higher (.0005) than the T group.

<i>Pre C</i>	<hr/>		<hr/>		<hr/>
	T	O	TE	E	I

On the R pretest scores the IMP Kit group was again significantly higher than the TE, O, and T groups (.0001); the E group, higher than the O and T groups (.0001); and the TE group, higher than the T group (.0005).

<i>Pre R</i>	<hr/>		<hr/>		<hr/>
	T	O	TE	E	I

The lack of any significant difference on the pretest scores between any of the pairs of experimental groups with respect to the amount of CR Gap is indicated by the appearance of all of the names of the groups under a common line.

<i>Pre C-R</i>	<hr/>				
	E	I	TE	T	O

On the posttest scores one less pair of groups was different on the C test; one more pair, different on the R test; and there emerged at this time a pair different with respect to the CR Gap. On the C posttest scores both the IMP Kit group and the E group were significantly higher (.0001) than the T and O groups, and the TE group was higher (.0005) than the T group as well as higher (.0001) than the O group. On the R posttest scores, the IMP Kit group was significantly higher (.0001) than every one of the other groups; the E group, higher (.0001) than the T and O groups; and the TE group, higher (.001) than the T group and higher (.0001) than the O group. With respect to the amount of CR Gap on posttest scores, the one pair significantly different resulted primarily from the large pretest-posttest improvement in the R test score of the IMP Kit group; thus the CR Gap of the IMP Kit group turned out to be significantly smaller (.0001) than that of the TE group.

On the improvement indicated by the difference between posttest and pretest scores, the IMP Kit group clearly emerges as the group that achieved the greatest improvement. The only other group that was significantly higher than any of the other groups on any of the three improvement measures was the TE group. In improvement on C test scores only one pair of groups was significantly different: the improvement of the TE group was greater (.005)

than that of the O group. In improvement on R test scores the IMP Kit group was significantly greater (.0001) than every one of the other groups, and the TE group was greater (.001) than the O group. With respect to what is probably the most important measure of all — the extent of the improvement in reducing the CR Gap — the IMP Kit group is the only group significantly better than any of the other groups. It deserves emphasis that the IMP Kit group turned out significantly better on this measure than every other group except the E group — and better than the E group, although not significantly so. The improvement of the IMP Kit group in reducing the CR Gap was greater (.0005) than that of the O and T groups and greater (.0001) than that of the TE group.

The significant differences on the C pretest and R pretest scores among the experimental groups deserve close scrutiny. Most (10 of the 13) of the differences are differences between EQUATIONS groups and nonEQUATIONS groups, and nearly half (6 of 13) are differences between the IMP Kit group and other groups. This raises the question as to whether the EQUATIONS groups generally and the IMP Kit group in particular were not simply more capable students at the beginning of the experiment. If so, perhaps it is not surprising that they improved more in learning to apply mathematical ideas during the two-week experiment. The next question is: Given that the students in the EQUATIONS groups were more capable at the end of their eighth-grade year when the experiment was conducted, were they also more capable two years earlier when they entered seventh grade?

School records indicate that the Stanford Arithmetic Test — Advanced (computation) was administered to all entering seventh-grade classes two years earlier and that 112 of the students in this study participated. From the scores recorded for this sample of the 237 students in the study for whom there is this indication of mathematical capability at the time of entry to the seventh grade, it appears that there was no significant difference between any of the pairs of the experimental groups at that time. In particular, there was no significant difference between the IMP Kit group and any of the other groups. Also, when data for the three EQUATIONS groups are combined and those for the two nonEQUATIONS groups are also combined, there is no significant difference in mean scores between the EQUATIONS groups and the nonEQUATIONS groups. The data are summarized in Table 1.

TABLE 1  
MATHEMATICAL CAPABILITIES TWO YEARS EARLIER: SCORES ON  
STANFORD ARITHMETIC TEST—ADVANCED (COMPUTATION)<sup>a</sup>

Group	All	I	E	TE	O	T	EQ	nEQ
N	112	22	14	25	23	28	61	51
$\bar{X}$	46.54	48.95	52.07	44.36	44.96	43.50	47.79	44.16
Sx	11.48	9.59	11.06	10.93	10.05	13.70	10.78	12.20

<sup>a</sup> Administered in September, 1972, for 112 of the 237 students in this study enrolled in Pelham Middle School eighth-grade classes in May, 1974.



The EQUATIONS and nonEQUATIONS groups were quite different two school years later when this experiment was undertaken, as was the IMP Kit group compared to all other groups except the E group. On the C pretest the EQUATIONS groups had a mean score of 6.38, significantly higher at the .0001 level than the 4.16 mean score of the nonEQUATIONS groups. On the R pretest the EQUATIONS groups were also significantly higher (.0001); the mean scores were 2.36 to 0.71. With respect to the CR Gap the 0.50 difference between the means of the two groups was not significant. The data for the EQUATIONS groups compared to the nonEQUATIONS groups are summarized in Table 2. The data for the other comparisons are in Figure 2.

TABLE 2  
COMPARISON OF EQUATIONS AND NONEQUATIONS GROUPS

	Pretest b		Posttest a		Posttest-Pretest a - b	
	EQ	nEQ	EQ	nEQ	EQ	nEQ
<u>C Test</u>						
N	87	119	84	118	84	104
$\bar{X}$	6.38	4.16	8.14	5.02	1.68	0.91
Sx	2.62	1.94	3.35	2.41	2.44	2.24
Signif.	.0001		.0001		.05	
<u>R Test</u>						
N	86	121	84	118	83	107
$\bar{X}$	2.36	0.71	4.18	1.18	1.72	0.50
Sx	1.94	1.08	3.17	1.55	2.20	1.19
Signif.	.0001		.0001		.0001	
<u>C-R</u>						
N	86	116	84	114	83	98
$\bar{X}$	4.02	3.52	3.96	3.91	-0.04	0.41
Sx	2.06	1.72	2.44	2.06	2.99	2.22
Signif.	ns (.0589)		ns		ns	

### Discussion

This study provides strong support for the proposition that skills in applying mathematical ideas can be improved by learning procedures that are rich in opportunities for such application at appropriate levels of complexity for each student. Interpreted most favourably, the results show that the combination of playing EQUATIONS over a two-year period and then working intensively with the IMP Kits for two weeks enables students to apply mathematical ideas (in the sense studied in this experiment) better than any of the other four sets of conditions do: better than just playing EQUATIONS alone, better than playing EQUATIONS and being taught explicitly by the teacher the 21 ideas presented in the IMP Kits, better than being taught the 21 ideas in an ordinary traditional mathematics class, and better than being in an ordinary traditional class without any special teaching of the ideas — and, furthermore, better in each case by a highly conservative test at an extreme level of significance (.0001). It should be acknowledged immediately that there are some questions with respect to this

most favourable interpretation which require further investigation. At the start of the experiment the IMP Kit group was clearly performing at a higher level of achievement than were the other groups. This superior performance seems linked to their two-year experience in playing EQUATIONS. The available evidence indicates that the EQUATIONS and nonEQUATIONS groups were not different upon their entry to the seventh grade, but that after two years of different experience with respect to whether or not they played EQUATIONS, the EQUATIONS group was significantly better in both computing with and applying these 21 mathematical ideas. The emphasis of the IMP Kit experience is clearly in the direction of improving skills in applying mathematical ideas, although it does improve both computing and applying. In terms of reducing the difference between understanding a mathematical idea in the sense of computing correctly with it and understanding it in the sense of being able to apply it in a context where it must be recognized to be relevant (the CR Gap), playing through the IMP Kits clearly is more effective than any of the other methods tried in this experiment (except possibly playing EQUATIONS alone, where the effect is in the right direction but is not significant).

For classrooms in which learning to apply mathematical ideas is still a problem, the implications of the findings of this study are obvious: learning environments so structured are effective — and should be used.

In an earlier study (Allen & Main, 1976), the objectives of designing structured learning environments of the type studied here were described as the enhancement of both the affective and cognitive dimensions. It was shown there that attitude as measured by absenteeism was profoundly affected through the use of such structured learning environments, and the prediction was ventured that in intelligent hands the achievement of knowledge should be, too. In this inner-city school in Detroit the learning structure we are designing has been in such hands. We still need to learn more, but the efforts there have advanced our understanding one solid step along the way.

We gratefully acknowledge the indispensable relevance of the cooperation of four dedicated educators at Pelham Middle School whose enthusiastic support made this study possible: Lewis Jeffries, Principal; Gloria Jackson, Mathematics Department Chairperson; and William Beeman and Harold Hauer, Mathematics Teachers.

#### *References*

- Allen, L. E. *EQUATIONS: The game of creative mathematics*. New Haven: Autotelic Instructional Materials Publishers, 1972. (Originally published, 1964.)
- Allen, L. E. RAG-PELT: Resource Allocation Games—Planned Environments for Learning and Thinking. *Simulation and Games*, 1972, 3, 407-438.
- Allen, L. E., & Main, D. The effect of instructional gaming upon absenteeism: The first step. *Journal for Research in Mathematics Education*, 1976, 7, 113-128.
- Allen, L. E., & Ross, J. Instructional gaming as a means to achieve skill in selecting ideas relevant for solving a problem. *International Journal of Mathematics Education and Science Technology*, 1975, 6, 475-488. The 21 published IMP Kits are available from the Autotelic Instructional Material Publishers, 1490 South Boulevard, Ann Arbor, MI. 48104 (\$15).
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw Hill, 1971.

J. P. BANA

*Mount Lawley College*

and

DOYAL NELSON

*University of Alberta*

## Some Effects of Distractions in Nonverbal Mathematical Problems

*The central purpose of this study was to investigate the effects of distractions in nonverbal problems on the problem-solving behaviour and performance of young children. Ten groups, matched on sex and ability, were selected from each of grades 1-3 in six schools. At each grade level, each group was randomly assigned to one of two partitive division problems in one of five problem settings representing different types and amounts of distraction. Two-thirds of the 360 subjects were distracted by irrelevant spatial-numerical or color-attribute cues. Both the problem setting and the degree of attention to distractions had a significant effect on problem-solving behaviour and performance. Distracted subjects appeared to identify and attempt a different problem from the one assigned. (Dr. Bana is Senior Lecturer in Mathematics Education at Mount Lawley College, Perth, Australia; Dr. Nelson is Professor in the Department of Elementary Education, The University of Alberta.)*

Problem solving has always been an important component of the school mathematics program. Although the research in this area has been extensive, it has been directed almost exclusively at verbal problems and still leaves many questions unanswered (Suydam, 1976). At the University of Alberta a research project was undertaken to study the behaviours of children aged 3 to 9 years in solving nonverbal problems (Nelson & Sawada, 1975). Some of the findings of this project have now been reported (Bourgeois, 1976; Little, 1976; Nelson & Kieren, 1976; Bourgeois & Nelson, 1977). As reported by Bourgeois and by Bourgeois and Nelson, the younger children were often distracted by irrelevancies in the problem setting.

Bana and Nelson (in press) undertook an exploratory study to establish some bases for a more detailed investigation of the role of distractions in nonverbal problems. In the exploratory study four groups of eight children, matched on sex and ability, were selected from each of grades 1-3 in a large



elementary school in Edmonton. At each grade level each group attempted a nonverbal partitive division problem embodied in one of four settings representing different types or amounts of distraction. It was found that both grade and problem setting had a significant effect on problem-solving performance. The extent of attention to distractions was dependent on the problem setting and significantly affected performance.

A number of studies have shown that the inclusion of irrelevancies in word problems increases problem difficulty (Biegen, 1971; Jerman & Rees, 1972; Blakenship & Lovitt, 1976). Irrelevant data also appears to hinder concept formation with pictorial material, as shown by Mansfield (1970). The Piagetian and related studies provide numerous instances where young children's centrations on irrelevant perceptual cues in concrete and pictorial materials prevent them from developing specific concepts. For example, children who failed to conserve numerosness often focused on the configuration of a set of objects and ignored the number of objects (Piaget, 1952). Gelman (1969) and Bryant (1974) found that preschool children succeeded in conserving number after being trained to ignore irrelevant cues. Although there is some evidence that young children are also distracted in nonverbal mathematical problems, no studies have systematically investigated the role of distractions in such problems. The major purposes of the current study were to determine the effects of distractions in nonverbal problems on the problem-solving behaviour and performance of young children in grades 1-3, and investigate the relationships of distractedness (whether a child was distracted or not) to sex, ability, conceptual tempo, grade, problem difficulty, and problem setting.

### *Method*

Two nonverbal partitive division problems of different difficulty levels were devised for the study, as follows:

1. In Problem A 12 cars were to be partitioned into three equivalent sets and the subject was required to give the number in each set.
2. Problem B was the same as Problem A except that there were 13 cars, which therefore involved a remainder.

Each of these problems was embodied in five problem settings representing different types and amounts of distraction, as follows:

1. Twelve (or 13) yellow toy cars of similar size and shape, together with three identical yellow toy trucks, each of which could take up to six cars (minimal distraction), as illustrated in Figure 1.
2. Cars and trucks as in (1) with a roadway for the trucks, a parking area for the cars, a building, a loading ramp, and three toy men — one near the ramp, and one on the back of each of the first two trucks in line (situational distraction), as in Figure 2.
3. A blue, a yellow, and a red truck, together with three blue, four yellow, and five (or six) red cars, all objects of the same shapes and sizes as before (color-attribute distraction), as in Figure 3.
4. Cars and trucks as in (1), but with six bays marked on the back of each truck (spatial-numerical distraction), as in Figure 4.
5. All the distractions of (2), (3), and (4) combined (maximum distraction), as in Figure 5.

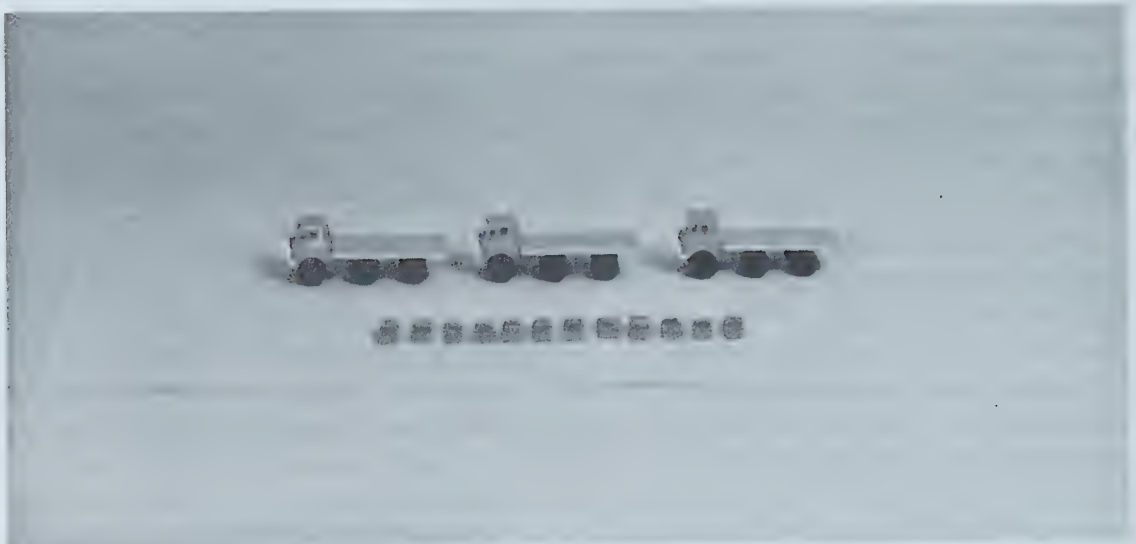


Figure 1. First problem setting.



Figure 2. Second problem setting.



Figure 3. Third problem setting.

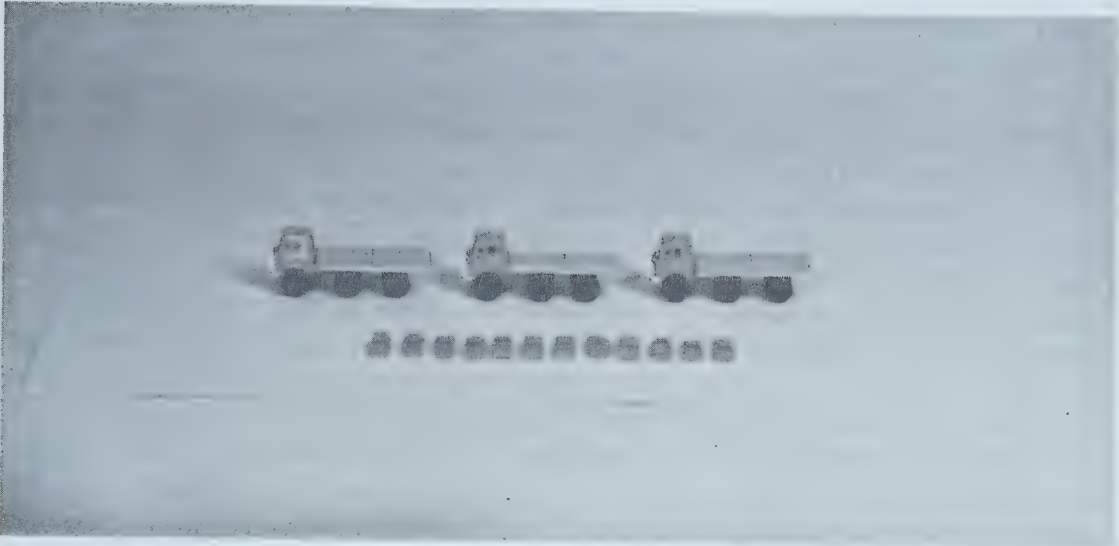


Figure 4. Fourth problem setting.

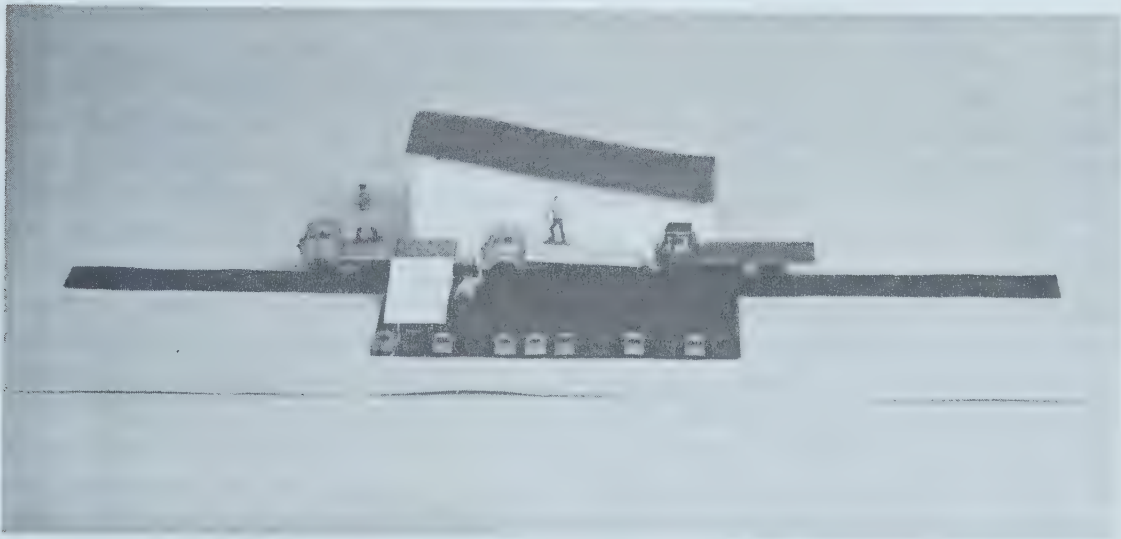


Figure 5. Fifth problem setting.

The types of distractions used here were based on the findings of Bourgeois (1976). In the exploratory study, Problem A was assigned to subjects in each of the first four settings (Bana & Nelson, in press).

The data for the main study were collected in November and December, 1976. Six schools were chosen from within the Edmonton Public School system so as to represent, as far as practicable, the population of first-, second-, and third-graders in the system. Stratified random sampling procedures, based on sex and Metropolitan Readiness Test percentile ranks, were used to select ten matched groups from the grade 1 pupils in these schools. Each group consisted of three girls and three boys in the high ability range, and three girls and three boys in the low ability range. Ten matched groups were selected from each of grades 2 and 3 in the same way, except that percentile ranks on a standardized mathematics test of the Edmonton Public School Board were used as the criterion for ability.

At each grade level each group was randomly assigned to either Problem



A or Problem B in one of the five settings described above. Each of the 360 subjects attempted the problem individually with the experimenter (*E*). The subject (*S*) was shown the materials, then *E* stated: "We have to load as many of those cars as we can onto the three trucks, but we must put the same number of cars on every truck. How many of those cars will be on each truck?" *E* repeated the problem question, and did so again if *S* was still unsure, asked any questions, or seemed uncertain how to respond. When *S* gave a solution *E* asked, "Are you sure that's right?" This question was repeated for each solution given, until *S* either confirmed a solution or indicated the unlikelihood of any further progress. Then *E* asked a question of the form, "How did you know there would be . . . on each truck?" Thereafter, the structure of the interview varied according to the subject's actions and verbalizations. The major objectives were to get at the thought processes used in attempting the problem, and to determine whether or not the subject was distracted. Verbalizations were recorded on audio-tape, and all observed nonverbal behaviors were noted on data cards. Samples of interview transcripts may be found in Bana (1977).

Each subject was awarded a score of either zero, one, two, or three on the problem-solving task. One point was awarded if the subject's first solution was correct. A second point was given if the subject confirmed the correct solution and demonstrated confidence in it, and a third point was allocated if a valid process was used in attempting the problem. The time taken for each subject to give the first response was noted and, based on a median split, was classified on the dimension of conceptual tempo as being either a fast (impulsive) or a slow (reflective) responder. A five-way analysis of variance was carried out to determine the main effects of grade, sex, ability, problem difficulty, and problem setting on problem-solving performance, and also any interaction effects. Chi-square goodness-of-fit tests, and tests for differences between means or proportions were used to investigate other relationships. An independent coder was employed to carry out a reliability check on the coded data. Based on a stratified random sample of 30 students, the coefficient of agreement was 0.94. A specific comparison of the scores awarded was also undertaken. Only two of the 30 scores differed from those of the experimenter, thus giving a reliability coefficient of 0.93. Based on these results, experimenter reliability was considered to be adequate.

### *Results*

The subjects exhibited a wide range of problem-solving behaviours which, in many cases, were influenced by the problem setting. Thirty-two percent of the subjects attempted the problem by loading cars on the trucks, while the remainder gave verbal responses without handling the material. Manipulations were less frequent in the fourth and fifth problem settings than in any of the other three settings. The tendency to manipulate the objects decreased only slightly through the three grades and the differences were not significant.

The more common arrangements of cars attempted, and the rationales behind these are given in Table 1. An arrangement such as 3,4,5 indicates that the subject said there would be, or actually put three cars on the first truck, four cars on the second, and five cars on the third truck in line. Only 36 percent of the subjects used a valid process in attempting the problem. Of

TABLE 1  
FREQUENCY DISTRIBUTION OF RATIONALES FOR SPECIFIC  
ARRANGEMENTS OF CARS

Rationale	Arrangement												
	3	3	3	4	5	5	5	6	6	6	9	Other Total	Other Total
	3	4	4	4	5	5	5	6	6	6	9		
	3	5	6	4	2	3	5	0	1	6	9		
Division or Multiplication	-	-	-	4	-	-	-	-	-	-	-	-	4
Distribution of Cars One by One	-	-	-	11	-	-	-	-	-	-	-	-	11
Trial of Equivalent Sets	36	-	-	109	7	3	4	1	1	-	-	24	185
Guess	3	-	-	3	-	-	2	-	-	1	1	2	12
Same Number of Cars as Trucks	6	-	-	-	-	-	-	-	-	-	-	-	6
Spatial-Numerical Cues:													
(a) Amount of Room	5	-	-	9	4	9	7	8	7	9	7	36	101
(b) The Marked Bays	-	-	-	-	2	2	10	10	6	87	-	12	129
Color-Attribute Cues:													
(a) Cars of Each Color	-	12	9	2	-	-	-	-	-	-	-	3	26
(b) Cars of a Particular Color	7	-	-	3	-	-	-	-	-	3	-	-	13
Induced Spatial-Numerical Cue	-	-	-	6	1	-	6	2	1	7	2	2	27
Induced Color-Attribute Cue	2	-	-	2	-	-	-	-	-	-	-	-	4
No Explanation	6	-	-	2	-	-	1	-	-	3	-	1	13
Other	-	-	-	-	-	-	-	-	-	-	-	3	3
Total	65	12	9	151	14	14	30	21	15	110	10	83	534*

\* Some subjects attempted more than one arrangement.

these, only 15 subjects employed a systematic process; the remainder used trial and error methods. As can be seen in Table 1, most of the subjects who used an invalid process based their solutions on irrelevant spatial-numerical cues. Color-attribute cues were also used as rationales for solution strategies. Some subjects stated that they obtained a solution by guessing, while others did not give any explanation for their particular arrangements of cars.

The frequencies of distracted subjects in each problem setting are given in Table 2. Spatial-numerical cues distracted 208 subjects. In the fourth and fifth settings the majority of the subjects said there would be six cars per truck because there were six bays marked on each truck. In the first three settings, where the trucks were unmarked, some subjects based their response on the amount of room on the trucks. However, the spatial-numerical distraction was not as strong in these cases. Manipulative solvers who were distracted by the amount of room or the number of bays would fill the first two trucks and put the remaining cars (if any) on the third truck. This accounted for the arrangements 5,5,2 or 6,6,0 for Problem A, and 5,5,3 or 6,6,1 for Problem B. The men on the first two trucks in the second and fifth settings distracted some subjects, but only in a spatial-numerical sense since each man was considered to occupy a space, which meant that one car fewer could be put on those trucks. The arrangements 3,4,5 for Problem A and 3,4,6 for Problem B were based on the numbers of blue, yellow, and red cars respectively, and thus occurred in the third and fifth settings. Some subjects based their arrangements on cars of a particular color; for example, seven subjects said there would be three cars on each truck because there were only three blue cars for the blue truck. There were 42 cases where subjects were distracted by

TABLE 2  
FREQUENCY DISTRIBUTION OF SUBJECTS DISTRACTED IN  
EACH PROBLEM SETTING

Problem Setting	Distraction			
	Spatial-Numerical	Color-Attribute	Other	Total*
1	28	—	—	28
2	36	—	—	36
3	18	32	—	46
4	64	—	—	64
5	62	10	1	64
1-5	208	42	1	238

\* Totals do not always correspond to row frequencies since some subjects attended to more than one type of distraction.

color-attribute cues and set out to match cars and trucks by color. Color was a stronger distraction than the spatial-numerical cue in the third setting. However, in the fifth problem setting, with the marked bays, the reverse was true.

Some subjects overcame distractions; that is, they were distracted by an irrelevant cue initially but then abandoned this as a means for solving the problem. Fifty-seven subjects overcame at least one distraction as shown in Table 3. The proportion of distracted subjects decreased with an increase in grade level, and the difference between observed and expected frequencies was significant ( $\chi^2 = 11.96, p < 0.005$ ). Subjects often gave no explanation for their attention to irrelevancies, particularly when they used a spatial-numerical cue. The most common reason given in this case was that there

TABLE 3  
OBSERVED AND EXPECTED FREQUENCIES OF DISTRACTED SUBJECTS  
IN EACH GRADE WHO SUCCEEDED OR FAILED IN  
OVERCOMING DISTRACTIONS

Grade	Overcame at Least One Distraction	Failed to Overcome Any Distractions	Total Distracted
1	12 (20.6)	74 (65.4)	86
2	19 (19.9)	64 (63.1)	83
3	26 (16.5)	43 (52.5)	69
1-3	57	181	238



were enough cars to fill all the trucks. For the subjects who were distracted by color, the prevalent explanation was in terms of aesthetics. When asked why they would match cars and trucks by color, children often replied that this would look “nice,” “pretty,” “good,” “better,” or “neat.” In some cases where a subject did not give an explanation for a particular response, a distraction was induced inadvertently by the experimenter’s questions. For example, if a subject said there would be six cars on each truck the experimenter may have asked: “Did you think about filling each truck?” If the subject replied in the affirmative this was recorded as an “induced” spatial-numerical distraction, but there was no way of interpreting the information gained in this way.

Distractedness has a very significant effect on problem-solving performance, as revealed by the difference between means ( $z = 9.69, p < 0.0002$ ). The mean score for the 238 distracted subjects was 0.58, while the mean for the remaining 122 subjects was 1.88. The mean scores for each grade on the two problems are given in Table 4. Scores were comparatively low in the fourth and fifth problem settings. A five-way analysis of variance revealed that ability, grade level, problem difficulty, and problem setting each had a very significant effect on problem-solving performance, as shown in Table 5. Although boys scored higher than girls in grade 1, the reverse was true in each of the other two grades and, overall, sex was not a significant factor. Girls performed better than boys in the third and fifth settings, where they seemed better able to handle the color-attribute distraction. The mean scores in Problem A and Problem B were 1.34 and 0.69 respectively. Only a small number of children were able to handle the problem involving a remainder, where the irrelevancies were more distracting than in Problem A. Many subjects tried to partition all 13 cars into three equivalent sets. The mean scores for the high and low ability groups were 1.19 and 0.84 respectively. Only those interaction effects with a significance level below 0.10 are reported in Table 5. Scores on Problem A were much the same for both ability groups, but scores for the low ability group fell sharply in Problem B, whereas the high ability group’s scores fell only slightly. First-graders scored low on both problems. However, performance on Problem A increased sharply in grade 2 and continued to increase in grade 3, while scores on Problem B increased very slowly from grade to grade.

TABLE 4  
MEAN SCORES ON BOTH PROBLEMS FOR EACH GRADE IN THE  
FIVE PROBLEM SETTINGS

Grade	Problem Setting					
	1	2	3	4	5	1-5
1	0.46	0.83	0.88	0.46	0.08	0.54
2	1.58	1.38	1.46	0.38	0.75	1.11
3	1.79	1.83	1.75	0.63	1.00	1.40
1-3	1.28	1.35	1.36	0.49	0.61	1.02

TABLE 5  
SELECTED RESULTS OF FIVE-WAY ANOVA

Source*	df	MS	F Ratio	Probability
A	1	2.844	2.349	0.12672
B	1	37.378	30.862	0.00000
C	1	11.378	9.395	0.00243
D	2	22.858	18.874	0.00000
E	4	13.357	11.029	0.00000
B $\times$ C	1	4.900	4.046	0.04541
A $\times$ D	2	2.953	2.438	0.08949
B $\times$ D	2	5.853	4.836	0.00876
B $\times$ C $\times$ E	4	2.768	2.286	0.06086
A $\times$ B $\times$ C $\times$ E	4	3.857	3.185	0.01422
Error	240	1.211		

\* A = Sex, B = Difficulty, C = Ability, D = Grade, E = Setting

The relationship between conceptual tempo and distractedness is illustrated in Figure 6. The mean scores for reflective and impulsive subjects were 1.42 and 0.61 respectively, with the difference being highly significant ( $z = 6.27$ ,  $p < 0.0002$ ). The nondistracted reflective group with a mean score of 2.20 was far superior to the distracted impulsive group which had little or no success with a mean scores of 0.21 ( $z = 13.49$ ,  $p < 0.0002$ ). Reflective subjects were no less distracted than impulsive subjects, but most of those who overcame distractions were in the reflective group. Conceptual tempo was independent of sex, ability, and grade level.

### *Discussion*

The findings indicate that young children are easily distracted by irrelevancies in a problem setting. It appears that distracted children identify and attempt a different problem from the one assigned. In the fourth and fifth settings, with bays marked on the trucks, most children seemed to think they had to find out how many cars would fit on each truck, so they simply counted the bays to solve their problem. In the other three settings children would often count spaces on the trucks or load cars to see how many would fit. Some children who were distracted by color appeared to see the problem as one requiring the matching of cars and trucks according to color. Others thought the required task was to match colors and, at the same time, put the same number of cars on each truck. The remainder involved in Problem B was not considered as a distraction for the purposes of this study, but it seemed to act as such. Many children tried to put all 13 cars on the three trucks and also get the same number on each.

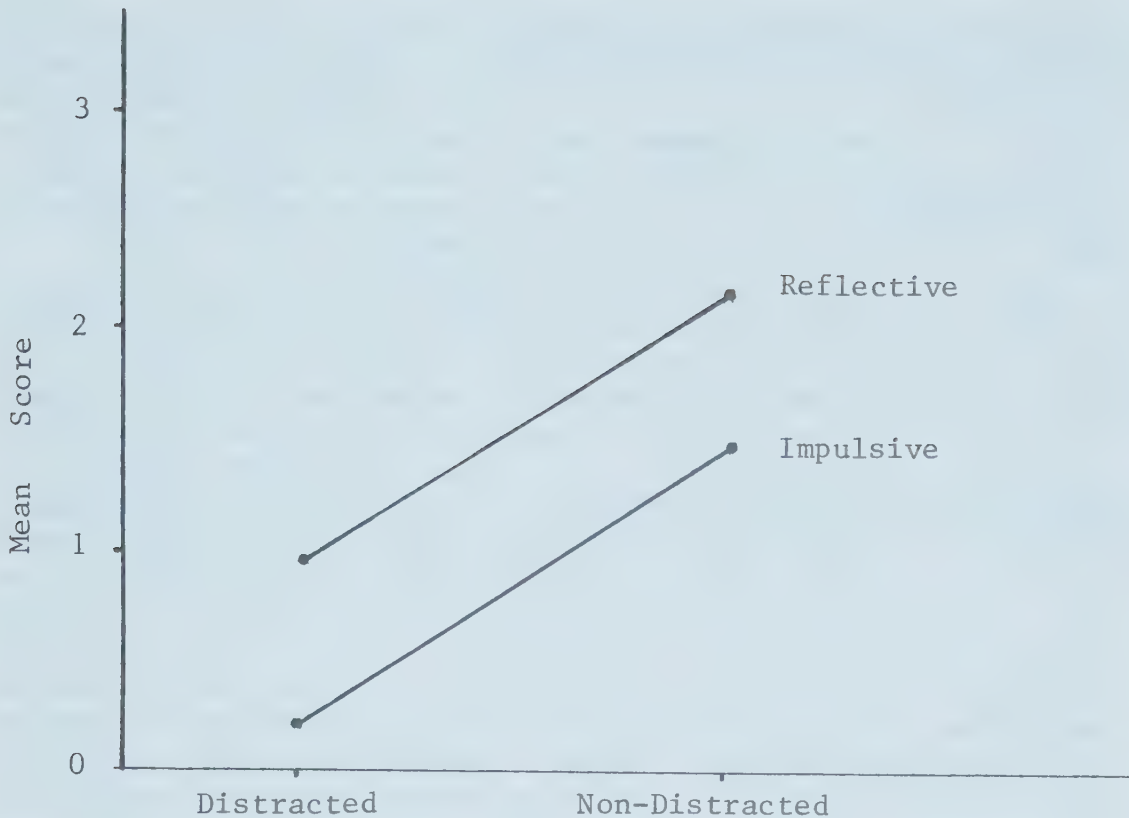


Figure 6. Relationship between distractedness and conceptual tempo.

For the distracted child the relevant information becomes irrelevant. For example, those who were distracted by spatial-numerical cues focused on the amount of room on the trucks and ignored the fact that there were either 12 or 13 cars to distribute. The situational distractions built into the second and fifth problem settings had no effect as such. Some children were distracted by the men on the trucks, but only in a spatial-numerical sense in that they were considered to take up room. Thus, it appears that a distraction only comes into play if it forms the basis for some plausible alternative problem for the child. Six of the subjects reasoned that there would be three cars on each truck because there were three trucks; thus apparently misinterpreting the problem question. It is likely that the verbal problem question will affect distractedness in the problem setting. The performance of subjects on Problem A in the exploratory study (Bana & Nelson, in press) was higher than in the main study and the difference may be due, at least in part, to the fact that the problem question was posed differently in each case. More research is needed to investigate the interaction of irrelevancies in the problem question with those in the problem setting.

Among verbal solvers, girls were less distracted than boys. Also, girls seemed better able to cope with color-attribute distractions. However, more research should be undertaken to establish any sex differences more definitively. With an increase in grade level fewer children were distracted, and those who were seemed better able to overcome the distraction. Developmental differences in problem-solving performance are apparently



related to the ability to cope with noise. This feature may be noted in the acquisition of Piagetian concepts. For example, children who can conserve numerosness have learned to ignore irrelevant perceptual cues such as length or configuration. However, as Suydam (1976) points out, different situations can yield different results on Piagetian tasks. Hence, whether a child is classified as a conserver or nonconserver probably depends on the effect of the irrelevancies in the particular setting.

Different settings can yield different performances. If a child succeeds in one setting teachers should not be deluded into thinking that the task has been mastered. Children should be subjected to the same problem in a variety of situations, along the lines of the multiple embodiment principle proposed by Dienes (1971). More research is needed to investigate the role of all the major types of distractions over a wide variety of problems. Skemp (1971) suggests that situations should be presented with few irrelevancies initially, then the noise level gradually increased. However, longitudinal studies would be needed to verify this. Children need to develop the ability to cut through noise in order to abstract mathematics from the environment. The findings suggest that young children are not given sufficient experience with manipulative materials or with partitive division problems, as indicated by the reluctance of many subjects to manipulate materials and the lack of systematic approaches. Finally, the low success rate of impulsive responders suggests that teachers should not hurry children's responses; rather, they should encourage children to reflect on each problem and thus provide more opportunity to cope with distractions.

#### References

- Bana, J. P. Distractions in nonverbal mathematical problems: Some effects on the problem-solving behavior and performance of young children. Unpublished doctoral thesis, University of Alberta, 1977.
- Bana, J. P., & Nelson, L. D. Distractors in nonverbal mathematical problems. *Journal for Research in Mathematics Education* (in press).
- Biegen, D. A. The effect of irrelevant and immaterial data on problem difficulty. Unpublished doctoral dissertation, University of Cincinnati, 1971.
- Blakenship, C. S., & Lovitt, T. C. Story problems: Merely confusing or downright befuddling? *Journal for Research in Mathematics Education*, 1976, 7, 290-298.
- Bourgeois, R. D. Young children's behavior in division problems. Unpublished doctoral thesis, University of Alberta, 1976.
- Bourgeois, R. D., & Nelson, L. D. Young children's behavior in solving division problems. *Alberta Journal of Educational Research*, 1977, 23 (3), 178-185.
- Bryant, P. E. *Perception and understanding of young children: An experimental approach*. London: Methuen, 1974.
- Dienes, Z. P. *Building up mathematics*. Fourth edition. London: Hutchinson Educational, 1971.
- Gelman, R. Conservation acquisition: A problem of learning to attend to relevant attributes. *Journal of Experimental Child Psychology*, 1969, 7, 167-187.
- Jerman, M., & Rees, R. Predicting the relative difficulty of verbal arithmetic problems. *Educational Studies in Mathematics*, 1972, 4, 306-323.
- Little, J. J. A mathematical and cognitive analysis of children's behavior in spatial problems. Unpublished doctoral thesis, University of Alberta, 1976.
- Mansfield, R. S. Developmental trends in the effects of noise on problem solving. Unpublished doctoral dissertation, Harvard University, 1970.

- Nelson, L. D. Problem solving in a model for early mathematics learning. In A. R. Osborne (Ed.), *Models for learning mathematics: Papers from a research workshop*. Columbus, Ohio: ERIC Centre for Science, Mathematics and Environmental Education, 1976.
- Nelson, L. D., & Kieren, T. E. Children's behavior in solving spatial problems. *Alberta Journal of Educational Research*, 1977, 23, 22-30.
- Nelson, L. D., & Kirkpatrick, J. Problem solving. In J. N. Payne (Ed.), *Mathematics learning in early childhood*. Thirty-seventh Yearbook of the National Council of Teachers of Mathematics. Reston, Va.: NCTM, 1975.
- Nelson, L. D., & Sawada, D. Studying problem solving behavior in young children: Some methodological considerations. *Alberta Journal of Educational Research*, 1975, 21, 28-38.
- Piaget, J. *The child's conception of number*. London: Routledge & Kegan Paul, 1952.
- Skemp, R. R. *The psychology of learning mathematics*. Harmondsworth, Middlesex: Penguin Books, 1971.
- Suydam, M. N. Research related to the mathematics learning process: 1970-1975. In *Forschung zum Prozess des Mathematiklernens*. Bielefeld: Institut für Didaktik der Mathematik der Universität Bielefeld, 1976.

LILLIAN WHYTE

*The University of Alberta*

## Logico-Mathematical and Spatial Development in Children Underachieving in Arithmetic

*The WISC, several Piagetian logico-mathematical and representational space tasks, the Frostig DTVP, and a test of motor ability were administered to elementary school children at three age and three arithmetic achievement levels, to determine whether specific patterns of cognitive and/or spatial development are related to arithmetic achievement, and whether the patterns vary with age.*

*Verbal intelligence as measured by the WISC emerged as a significant factor and at the three age levels; however, WISC performance IQ did not. WISC Digit Span and Arithmetic, reported in the literature to reflect a factor described as "freedom from distractibility," differentiated achievers from those severely deficient in quantitative thinking ability. Four Piagetian tasks, seriation, inclusion of classes, and length and number conservation, accounted for 71.69% of the variance at the seven-year level. At the nine-year level, WISC Vocabulary and a Piagetian measurement task accounted for 61.16 and 15.19% of the variance respectively. Figure Ground and Form Constancy from the Frostig were significantly related, but only at ages seven and nine. Motor ability was not significantly related. One of the five representational space tasks, the concept of opposition of left and right, was significantly related to arithmetic achievement but only at the eleven-year level, contributing 34.40% of the variance, while WISC verbal IQ contributed 31.26%.*

*It was concluded that specific patterns of cognitive and spatial-motor development do characterize elementary school children at different achievement levels and the patterns do vary with chronological age. (Dr. Whyte is Professor in the Department of Educational Psychology at The University of Alberta.)*

The teaching of mathematics in elementary grades traditionally has consisted in imparting content thought to be suitable for the majority of children at particular grade levels. Recently there has been emphasis on theoretical models of development and learning as the basis for instruction. Both traditional and modern approaches to curriculum development and instruction may be too narrow in scope. Often the child's emotional and cognitive development as well as specific cognitive deficits, factors which



could contribute to difficulty in learning mathematics content, are ignored. Furthermore, the tendency is to treat mathematics content as an entity, with little thought given to prerequisite skills which may vary as content varies.

The present study was designed to determine whether children at three age levels and three achievement levels in elementary school arithmetic differ in skills thought to be prerequisite to acquiring an understanding of the natural number system.

### *Related Theory and Research*

#### *The Nature of Mathematics Content*

There is little agreement in the literature concerning the factors which comprise mathematics content. Hyde (1970) and Lovell (1961, 1966) emphasized the importance of perception to the development of number. Vernon (1950) concluded from a review of research that while g-saturation (general intelligence) varies for various forms of arithmetic, there is general agreement in the factorial literature that one of the fundamentals of arithmetic is the general factor described as g.

Piaget's theory of number (1941) postulates a conceptual basis. Number, according to Piaget, is at the same time both class and asymmetrical relation, derives from the union of these logical operations, and is the result of the fusion of several related notions: conservation of continuous and discontinuous quantities; provoked correspondence and equivalence of corresponding sets; spontaneous correspondence and the cardinality value of sets; seriation; the synthesis of cardinality and ordination. These notions grow out of a pre-numerical and pre-logical period, are organized in an invariant sequence of stages, and culminate in a structure Piaget terms conservation, that is, "A set or collection (which) remains unchanged irrespective of the changes occurring in the relationship between the elements" (Piaget, 1941, p. 1).

The few studies attempting to verify that number results from the union of class and asymmetrical relation have not supported the hypothesis, e.g., Dodwell (1962). There is, however, a considerable body of research confirming a relationship between basic number concepts postulated by Piaget and arithmetic achievement, e.g., Dodwell (1961, 1962), Hood (1962), Beard (1963), Brown (1969). While operativity as defined by Piaget and general intelligence as defined by Vernon are both considered aspects of conceptual development, factor analytic findings tend to separate psychometric tasks and Piagetian tasks on different factors (Lunzer, 1971; Kohlberg & DeVries, 1974).

An alternative but controversial hypothesis also advanced is the notion that space is a fundamental factor in arithmetic content. While some researchers consider space an essential factor in arithmetic content, e.g., McFarlane-Smith (1964), others such as Vernon (1950) concluded k (space) and n (number) tests are strongly opposed, apart from g. Recent research in spatial development holds promise for resolving the controversy. Generated largely from the work of Piaget (1960, 1967), the research attempts to define and clarify the nature of spatial development.

Unfortunately, there is no extensive literature on the Piagetian tradition on arithmetic and spatial concepts to review; however, three studies are relevant to the present discussion. Mycock (1969) concluded the SRA Space Test had no high correlations with any of the Piagetian spatial tests but he

reported a spatial factor of a specific kind involving perception of similarities in geometrical figures. Goldschmid and Bentler (1968) administered number conservation tasks, two-dimensional space tasks, and three tasks on conservation of length to children from 5 to 7 years. They concluded the number conservation tests were not measuring exactly the same dimension as the area and length tasks. Conservation of number correlated .52 with arithmetic while length and area correlated .39. Goldschmid (1967) reported conservation of number tests were positively correlated with IQ, MA, and verbal ability. Both two- and three-dimensional space tasks correlated significantly with conservation tasks, although the range of correlations for the latter was much lower (.16 to .26) than for the former (.31 to .61). The correlations with total conservation were .66 and .35 for two- and three-dimensional space tasks respectively and .54 and .17 between number and two- and three-dimensional space tasks respectively. It would appear that two- and three-dimensional space tasks are fairly heterogeneous abilities.

The lack of consensus among researchers concerning the relationship of space and arithmetic is due in part to the failure of researchers to agree on what is meant by space. Dodwell (1971) stated the problem, "there is a valid distinction to be drawn between the apprehension, or discrimination, of forms and objects, and the understanding of their nature—or the conception of space and spatial relations. The weaknesses in modern perceptual theory come about largely through failure to observe this distinction" (p. 119).

Piaget (1967) does make the distinction, that is, a distinction between a perceptual space and a cognitive space. Cognitive space is operative and operational knowledge is dependent upon intelligence at all levels. Perceptual space is figurative and is dependent upon perception. Cognitive space becomes representational when the symbolic function appears around two years of age and, unlike perceptual space, develops in sequential stages.

These notions of space may have potential value for understanding the literature relative to space and arithmetic, particularly as it relates to learning disabilities. Research in learning disabilities appears to be concerned with perceptual space and arithmetic disability, emphasizing disabilities in motor development and visual perception as they relate to failure in learning arithmetic, e.g., Ayres (1965, 1975). The tendency for spatial disabilities and arithmetic disability to occur together is well documented in the literature. Not so well documented is the nature of the relationship.

### *Causes of Arithmetic Disability*

The etiology of arithmetic disability seems as varied as the factors which constitute arithmetic content. Lowered intellectual capacity is one cause of diminished ability to learn arithmetic but many intellectually capable children seem unable to acquire an understanding of number systems.

Although findings vary with grade and/or C.A. level, Hood (1962), Almy (1966), Goldschmid (1968), Reimer (1968), Wheatley (1970) and Riggs (1970) in general reported significant correlations between arithmetic achievement and ability to conserve. Results are not as positive when older children are assessed; for example Overholt (1964) found no significant differences in arithmetic achievement between fourth-grade conservers and nonconservers



when he controlled for IQ. While nonconservation in number concepts appears an etiological factor in arithmetic disability, at least in kindergarten and grade one, there is insufficient evidence concerning stage of development in classification tasks, largely because the research has not been done.

Spatial deficit is also postulated as an etiological factor in arithmetic disability. Blackall (1972) studied conservation, perceptual efficiency and estimation of length, area and numerical quantity in grade three children performing at three achievement levels in arithmetic. The three groups differed significantly in visual and auditory discrimination, memory and closure (all said to have a spatial theme and probably measures of perceptual space). Four of the five factors reported are relevant; Factor I contained arithmetic, intelligence, reading, visual discrimination and auditory memory and closure. Factor II was a conservation factor, Factor III was tactual in nature, while Factor IV was a representation factor, including tasks requiring drawings of shapes presented visually and tactually. It would appear from this study that both perceptual and representational space as well as conservation and general intelligence play a role in the degree of arithmetic achievement attained by grade III children.

There is a particularly severe form of arithmetic disability termed *acalculia* or *dyscalculia*, generally defined as a disturbance or impairment in the ability to do simple arithmetic, usually associated with neurological dysfunction or brain damage, and which interferes with quantitative thinking. The characteristics most often cited are deficiency in visual-spatial organization and nonverbal integration manifested by disturbance in ability to distinguish shapes, sizes, etc; disturbances in body image; apraxia; confusion of spatial relationships such as up-down, front-back; clumsiness in manual, postural and ambulatory operations (Johnson & Myklebust, 1967; Kaliski, 1961; Cohn, 1968). Arithmetic errors include malformed or large number symbols, inability to recognize operator signs or linear separators, and inappropriate ordering of numbers in multiplication.

Koppitz (1971), in a five-year longitudinal study, reported that children who had to be hospitalized manifested the lowest arithmetic achievement and the most extreme, aggressive, explosive and delinquent behaviour which she attributed to their exceptionally poor integration as measured by the Bender Gestalt Test. Research has shown a close correlation between performance on the Bender and arithmetic achievement, and both are related to children's integrative ability and their overall behaviour.

Finger agnosia, *agraphia* and left-right confusion, alone or in combination, are frequently found among children with *dyscalculia*. Werner and Carrison (1963) concluded that finger agnosia, a failure to recognize, show, and name the fingers of either hand, may be associated with a tactile dysfunction, a deficit in visual form and perception, or a disturbance in language function. Benton (1959) reported finger agnosia and right-left disorientation tend to occur together with the highest correlations (.62) between left-right orientation and finger localization, and calculation.

A developmental variant of constructional apraxia described by Ayres (1975) as a disturbance in motor planning and by Grewel (1951) as a disturbance of spatial handling, has been reported by Kinsbourne and Warrington (1963) and Ayres (1965). In the former study, subjects were best



represented by a marked specific disability in certain "spatial tasks," low WISC performance IQ and marked discrepancy between the subtest pairs of Block Design and Object Assembly, and Vocabulary and Digit Span, while Ayres reported the factor loading significantly on every test of tactile perception and apparently related to eye-pursuit and two aspects of body scheme: tactile impulses and integration of both sides of the body. Ayres also reported a substantial relationship between skill in number concepts and perceptual-motor functions.

Finally, Sabatino and Hayden (1970) in a study of 472 children failing elementary grades reported that many of the perceptual tasks were correlated significantly with WRAT arithmetic achievement scores for children less than C.A. 9-6 but there were considerably fewer significant correlations for children older than 9-6. They concluded that basic perceptual behaviours are not directly related to academic achievement for older children.

In summary, it is hypothesized that two etiologies of arithmetic disability are associated with cognitive and spatial disability. While low intellectual capacity is one cause, a major etiology appears to be associated with lack of operational level thought in logico-mathematical concepts. A second major etiology appears to be associated with a complex syndrome of spatial-motor disability. Motor and perceptual deficits which in turn may be caused by disorders in such systems as tactile, appear to be related to, or perhaps even form the basis of the problem. It is not clear whether the spatial problem is perceptual or conceptual in nature (or a complex combination of the two) but the greatest number of studies seem to have been concerned with perceptual space. The factor of age also complicates the problem, since prerequisite skills for arithmetic learning at younger ages do not appear to be the same for older children.

### *The Study*

It was hypothesized that two major causes of arithmetic disability are related to dysfunction(s) in cognitive and spatial-motor development. The present study compared the performance of elementary school children at three age levels and three achievement levels in arithmetic, on logico-mathematical and spatial development tasks.

### *Subjects*

Nine groups of children ( $N = 87$ ) at three C.A. levels (7, 9 and 11 years) and performing at three levels in arithmetic achievement (achievers, underachievers, and nonachievers) were included. Achievers were defined as children achieving from 75 to 100% of grade level expectation on a diagnostic arithmetic test covering content from grades I to VI, while underachievers and nonachievers achieved from 50 to 74% and less than 50%, respectively.

Achievers and underachievers were selected from a large elementary school in Edmonton, which drew its population from low, middle and high SES families. Nonachievers at ages nine and eleven were selected from adaptation classes designed for children with learning disabilities, while seven-year-old nonachievers were referred by school psychologists from two elementary schools. C.A. and IQ descriptions are presented in Table 1.

TABLE 1  
C.A. and IQ DESCRIPTION OF SUBJECTS

Group	N	C.A.			IQ								
		$\bar{X}$	SD	Range	WISC VERBAL			WISC PERFORMANCE			WISC FULL SCALE		
					$\bar{X}$	SD	Range	$\bar{X}$	SD	Range	$\bar{X}$	SD	Range
VII A	9	89	3.00	84-93	124.1	18.89	103-153	115.9	11.45	93-135	122.1	14.39	99-145
VII U	9	90.11	3.95	85-95	98.8	9.54	84-113	109.7	13.91	81-132	104.3	6.14	90-109
VII N	10	91.90	5.24	84-97	94.5	12.89	75-116	97.7	15.28	65-118	95.7	12.36	67-119
IX A	10	113.60	2.27	110-117	116.6	12.08	97-135	112.0	17.76	86-145	116.1	14.26	99-144
IX U	10	118.40	5.44	112-130	106.2	11.36	87-123	104.4	17.40	89-149	105.9	12.41	88-138
IX N	9	112.44	4.72	105-119	89.3	10.26	75-104	92.2	8.36	79-100	89.7	7.17	79-99
XI A	10	138.10	4.04	132-143	102.1	8.28	91-113	103.2	14.73	83-125	104.0	10.63	90-119
XI U	10	138.30	2.58	135-143	97.5	10.35	85-115	103.8	11.77	86-125	100.5	6.14	90-109
XI N	10	145.10	11.88	127-164	79.5	11.46	62-96	91.0	16.85	58-111	83.8	11.78	62-98

### Procedure

**Tests Administered.** The diagnostic arithmetic test was administered individually to each subject to determine achievement level and each was then administered a battery of tests. Three types of tests were administered: tests of cognitive development, spatial development and the diagnostic arithmetic test.

The WISC was administered to assess intelligence but performance on the subtests was also examined to determine whether specific patterns were associated with achievement and/or C.A. level.

Three types of Piagetian logical thinking tests were administered:

1. Inclusion and multiplication of classes, both classification tests;
2. A number conservation and a seriation test to assess number concepts;
3. Conservation and measurement of length, tests of infralogical development.

Three areas of spatial-motor development were assessed:

1. Visual perception: the five subtests of the Frostig Developmental Test of Visual Perception (1966) were administered, Eye-motor Coordination, Figure Ground, Form Constancy, Position in Space and Spatial Relations.
2. Motor development was assessed with the Oseretzky Test of Motor Development (Stott, 2nd revision, 1966). Five areas of motor ability are included: balance of the whole body, control of the upper body, locomotor control, manual dexterity, speed and precision and a total motor disability score.
3. Five Piagetian tests of representational space described by Laurendeau and Pinard (1970) were administered: (a) Stereognostic recognition of shapes, (b) Construction of a projective straight line, (c) Localization of topographical positions, and (d) Coordination of perspectives.

The diagnostic arithmetic test was designed for the study by the author with the assistance of Dr. Kenneth Lovell, University of Leeds, and included four subtests: counting concepts, place value concepts, properties or laws of



the natural number system and computational skills. Content was a representative sampling of the arithmetic curriculum from grades I to VI. The test was administered to subjects individually, scores for each section were qualified for each grade level, a total score from the combined subtests was computed, and the subjects were categorized as achievers, underachievers, or nonachievers by the total score earned.

### *Analysis of Data*

1. Means for the nine groups for each of 35 variables were computed and the Tukey Gap Value Statistic (Tukey, 1949) applied to detect significant differences between the means at  $p = .01$  level.
2. A Multiple Stepwise Regression Analysis was computed for the three chronological age levels to determine which variables were contributing the greatest percentages of variance for the variable, Total Arithmetic Score.

### *Questions Investigated*

Four questions were investigated:

1. Will elementary school children classified as achievers, underachievers, and nonachievers in arithmetic be characterized by specific patterns of development on the WISC, tests of logico-mathematical concepts, and/or tests of spatial development?
2. If specific patterns of development do characterize children at the three achievement levels, will the patterns also vary across the variable, chronological age?
3. Are one or more measures of spatial development related to arithmetic achievement?
4. Which of the four areas of the natural number system assessed are most seriously deficient at each of the three chronological age levels?

## *Results*

### *Cognitive Development*

#### *A. WISC:*

Group means for verbal and performance IQs have been presented in Table 1. WISC comparisons significant at  $p = .01$  level are presented in Table 2. For brevity in reporting results, Roman numerals VII, IX, and XI are used to indicate age levels and combined with abbreviations A, U and N for achievers, underachievers, and nonachievers respectively to indicate achievement levels.

The seven achievers (VII A) were an intellectually superior group and significant differences were found for verbal IQ between VII A and every group except IX A. In performance IQ, however, VII A differed significantly with only two groups, IX N and XI N. For verbal IQ there appeared to be a pattern. Both achievers and underachievers differed significantly from nonachievers. In only one case (VII A and VII U) did achievers differ significantly from underachievers. The pattern was not repeated for performance IQ, where only one significant comparison was found, between IX A and IX N. All other significant comparisons were across age groups.



TABLE 2  
DIFFERENCES IN WISC MEAN SCORES

Group	WISC Variable	Tukey Gap Values							
		VII U.	VII N	IX A	IX U	IX N	XI A	XI U	XI N
VII A	Verbal IQ <sup>a</sup>	25.3	29.6		17.9	24.5	22.0	26.6	44.6
	Performance IQ <sup>b</sup>					23.9			24.9
	Information <sup>c</sup>	4.6	4.8			5.2	3.6	4.3	7.1
	Arithmetic <sup>d</sup>	3.9	4.6		3.8	6.2		4.4	5.4
	Vocabulary <sup>e</sup>	3.7	4.5			5.0	3.7	3.9	7.7
	Digit Span <sup>f</sup>	4.2	3.3		3.3	5.3		3.8	4.9
	Coding <sup>g</sup>					3.7			4.5
VII U	Verbal IQ			17.8					19.3
	Performance IQ								18.7
	Vocabulary			2.9					4.1
	Digit Span						3.1		
VII N	Coding								3.97
	Verbal IQ			22.1					
	Arithmetic			3.9					
	Vocabulary			3.8					3.2
IX A	Coding			3.9					
	Verbal IQ					27.0		19.1	37.1
	Performance IQ					19.8			21.0
	Information					3.2			5.1
	Arithmetic					5.5		3.7	4.7
	Vocabulary					5.3	3.0	3.2	7.0
	Digit Span					3.0			2.6
IX U	Coding					5.8		4.0	6.7
	Verbal IQ					16.6			26.7
	Information								4.3
	Vocabulary					3.1			5.8
	Digit Span						4.2		
	Coding								3.8
	XI A	Verbal IQ							
Information									3.5
Vocabulary									4.0
Digit Span								2.7	3.8
XI U	Verbal IQ								18.0
	Vocabulary								3.8

Tukey Gap Values significant at .01 level: <sup>a</sup>15.17; <sup>b</sup>18.50; <sup>c</sup>3.208; <sup>d</sup>3.723; <sup>e</sup>2.947; <sup>f</sup>2.640; <sup>g</sup>3.547.

There were no significant comparisons for the Picture Completion subtest and only three, two, and one for Picture Arrangement, Block Design, and Object Assembly respectively. Significant comparisons for the remaining subtests are reported in Table 2. Vocabulary was the subtest most often involved in the comparisons found to be significant, with 18 of 36 possible comparisons being significant, followed by Digit Span with 12 and Arithmetic and Information each with eight. Except at the seven-year level where VII A differed from both VII U and VII N, the largest number of significant comparisons within age levels were between achievers and nonachievers with very few significant comparisons between achievers and underachievers or underachievers and nonachievers. The remain significant comparisons were across age levels.

B. *Logico-Mathematical Tests:*

The mean stage of development for each of the Piagetian tests is reported in Table 3, while comparisons which were significant are reported in Table 4. To analyze the data, the stages and substages were converted to a numerical scale from 1 to 7 as follows:

Stage	Period of Development	New Scale
0	No response, presumably sensorimotor	1
1A		2
1B		3
2A	Preoperational Period	4
2B		5
3A		6
3B	Concrete Operational Period	7

*Classification Tests.* Three groups of achievers were functioning at the concrete operational stage in inclusion of classes, while underachievers and nonachievers with one exception (IX U) were preoperational. The nine- and eleven-year achievers and underachievers were operational in multiplication of classes. While differences at each age level were slight, there was a tendency for nonachievers to be less advanced than the two other groups.

*Number Concepts' Tests.* Except for groups VII A and VII N, most subjects were operational in number conservation. Only two groups, VII A and IX A, were functioning at the expected stage on the seriation task. Normal children tend to become operational on this task at approximately nine years. Several children in group XI A in the low normal range of intelligence on the WISC were preoperational on the seriation task.

*Infralogical Tests.* Most children were concrete operational on the

TABLE 3  
STAGE OF DEVELOPMENT ON LOGICO-MATHEMATICAL AND REPRESENTATIONAL SPACE TASKS

Test	Stage of Development								
	VII A	VII U	VII N	IX A	IX U	IX N	XI A	XI U	XI N
<u>Classification Tests</u>									
Inclusion of Classes	7.0	3.89	4.20	6.80	6.80	5.89	6.25	5.40	5.40
Multiplication of Classes	5.89	4.33	4.45	6.60	6.00	5.28	6.45	6.60	5.60
<u>Number Concept Tests</u>									
Conservation of Number	7.0	5.44	5.05	7.0	7.0	6.39	7.0	7.0	7.0
Seriation	5.22	3.56	3.20	6.20	4.45	4.44	5.70	5.40	4.60
<u>Infralogical Tests</u>									
Conservation of Length	6.56	5.89	6.20	7.0	7.0	6.33	7.0	7.0	6.80
Measurement	4.17	2.61	2.60	6.40	5.60	3.44	7.0	6.80	5.20
<u>Representational Space Tests</u>									
Stereognostic Recognition of Shapes	6.33	6.22	6.20	6.70	6.50	6.00	6.70	6.50	6.00
Construction of a Projective Straight Line	5.00	5.22	5.20	6.70	6.60	5.56	6.90	6.10	5.90
Localization of Topographical Positions	6.67	6.11	6.10	6.70	6.10	5.00	6.30	6.40	6.10
Concept of Opposition of Left and Right	5.00	3.00	3.70	4.20	3.70	3.78	5.60	4.80	3.80
Coordination of Perspectives	3.67	3.11	3.20	5.00	4.80	4.11	4.80	5.20	4.60

TABLE 4  
DIFFERENCES IN LOGICO-MATHEMATICAL AND REPRESENTATIONAL  
SPACE TEST MEAN SCORES

Group	Test	Tukey Gap Values							
		VII U	VII N	IX A	IX U	IX N	XI A	XI U	XI N
VII A	Inclusion <sup>a</sup>	3.12	2.80					1.60	1.60
	Multiplication <sup>b</sup>	1.60							
	Number <sup>c</sup>	2.56	1.95						
	Seriation <sup>d</sup>		2.02						
	Measurement <sup>f</sup>			2.24			2.84	2.64	
	Straight Line <sup>g</sup>			1.70	1.60		1.90		
	Topographical Positions <sup>h</sup>					1.66			
	Concepts of Left & Right <sup>i</sup>	2.00							
VII U	Inclusion			2.92	2.92	2.00	2.32	1.58	1.58
	Multiplication			2.27	2.27		2.12	2.27	
	Number			1.56	1.56		1.56	1.56	1.56
	Seriation			2.65			2.10	1.85	
	Length <sup>e</sup>			1.12	1.12		1.12	1.12	
	Measurement			3.79	2.99		4.39	4.19	2.59
	Straight Line			1.48			1.60		
	Concepts of Left & Right						2.60		
	Coordination of Perspectives <sup>j</sup>			1.89				2.09	
VII N	Inclusion			2.60	2.60	1.68	2.00		
	Multiplication			2.15	2.15		2.00	2.15	
	Number			1.95	1.95	1.33	1.95	1.95	1.95
	Seriation			3.00			2.45	2.20	
	Measurement			3.80	3.00		4.20	4.20	2.60
	Straight Line			1.50			1.70		
	Coordination of Perspectives							2.00	
IX A	Seriation				1.76	1.76			
	Measurement					2.56			
	Topographical Positions					1.70			
IX U	Measurement					2.16			
IX N	Measurement						3.56	3.36	
	Topographical Positions						1.30	1.40	
XI A	Measurement								1.80

Tukey Gap Values significant at .01 level: <sup>a</sup>1.588; <sup>b</sup>1.588; <sup>c</sup>1.61; <sup>d</sup>1.761; <sup>e</sup>1.039; <sup>f</sup>1.792; <sup>g</sup>1.484; <sup>h</sup>1.133; <sup>i</sup>1.944; <sup>j</sup>1.825.

conservation of length task and the majority of children in groups IX A, XI A and XI U were also operational on the measurement task.

In most instances the trend was increased average stage of development as C.A. increased, e.g., for nonachievers mean stages on the measurement task were 2.60, 3.44 and 5.20 for seven-, nine-, and eleven-year-olds respectively. At each age level average stage was lower as achievement level decreased, e.g., in measurement the mean stages for seven-, nine-, and eleven-year-olds respectively were 4.17, 6.40, 7.0 for achievers, 2.61, 5.60 and 6.80 for underachievers and 2.60, 3.44 and 5.20 for nonachievers. While there were differences, they were not always significant and the majority were between age levels as might be expected (Table 4). At the seven-year level the



differences were between achievers, and underachievers and nonachievers, but not between underachievers and nonachievers. Except for measurement the same trend was found at nine years, while at eleven years one significant comparison was found (XI A and XI N).

The number of tasks where comparisons were significant decreased as C.A. increased, that is, significant comparisons were found for six, two, and one tasks for seven-, nine-, and eleven-year-olds respectively. While logico-mathematical concepts appear to be prerequisite for learning arithmetic, development of operational thought at older ages may not result in academic achievement. Possibly other complicating factors such as motivation are involved, or it may be that operational thought, while essential, is not sufficient for learning arithmetic.

### *Spatial-Motor Development*

#### *A. Frostig Developmental Test of Visual Perception:*

No significant comparisons were found for the Eye-Motor Coordination and Spatial Relations subtests. Fourteen comparisons were significant, but only four were within age levels and three involved Form Constancy. Group VII A differed from VII N on Figure Ground and Form Constancy (Tukey Gap Values 15.6 and 23.8 respectively). Groups IX A and IX U differed significantly from IX N on Form Constancy (Tukey Gap Values 22.8 and 24.6). The remaining significant comparisons were across age levels.

#### *B. Oseretzky Test of Motor Ability:*

Except for manual dexterity, speed and precision and total motor error score, performance on this test was almost error free. Two comparisons were significant for the manual dexterity subtest, groups VII U and XI N, and IX and XI N. Two comparisons involving the speed and precision subtest were significant between groups VII A and XI N, and XIA and XI N. Four comparisons for total motor score were significant, group VII A, VII U, IX A and IX U, when compared with XI N. There was a tendency at each age level for the average error score to increase as the degree of arithmetic disability increased.

#### *C. Representational Space Tests:*

Mean stages on the five space tasks are reported in Table 3, while comparisons which were significant are reported in Table 4. All subjects were operational on the stereognostic recognition of shapes test. Subjects at the seven-year level and groups IX N and XI N were preoperational on the task, construction of a projective straight line. Seven comparisons for the task were significant but all were across age levels. One group, IX N at stage 2B, was preoperational on the localization of topographical positions test but four comparisons involving the task were significant; however, only one comparison (between IX A and XI N) was within an age level.

No group was operational on either the concept of opposition of left and right test or the coordination of perspectives test; however, Laurendeau and Pinard (1970) reported that normal children range from C.A. ten to twelve before operational thought appears for the former task, while stage 3 generally is not initiated for the latter task before 13 years. For the former task average stages at the seven-year level were consistent with norms (Laurendeau & Pinard, 1970) but at nine years, groups were slightly below

expected stages. At eleven years most subjects should be operational but only group XI A subjects were. Subjects were functioning at or slightly above expected levels on the coordination of perspectives test. Differences between groups were minimal. The concepts of left and right test differentiated seven-year achievers and underachievers, while nine-year achievers and nonachievers differed significantly on the localization of topographical positions test. All other significant comparisons were across age levels.

Arithmetic Achievement

Although subjects were categorized by total score from the combined arithmetic subtests, performance on each subtest was analyzed separately. Means are presented in table form (Table 5). For brevity, significant comparisons are not reported in table form but are available from the author.

TABLE 5  
AVERAGE SCORES ON DIAGNOSTIC ARITHMETIC TESTS

Group	Average Arithmetic Scores				Total Score
	Counting	Place Value	Properties	Computational Skills	
VII A	15.00	20.56	6.11	5.67	47.33
VII U	12.33	10.89	3.56	1.56	28.33
VII N	9.70	3.20	3.10	1.80	17.60
IX A	23.00	22.50	10.30	32.10	87.90
IX U	20.40	22.50	8.30	21.20	72.60
IX N	13.22	9.22	2.33	9.22	34.00
XI A	25.20	24.70	14.30	44.70	108.90
XI U	23.00	23.90	11.40	33.90	92.20
XI N	16.30	15.50	2.90	16.40	51.10

All groups different significantly on total arithmetic score, both within and across age levels. At the seven-year level achievers differed significantly from underachievers in place value concepts, while achievers and underachievers at nine and eleven years also differed significantly in one area, computational skills. Achievers differed significantly from nonachievers in every area. At seven years underachievers differed from nonachievers in one area, place value, while IX U and IX N, and XI U and XI N differed significantly in every area. Other significant comparisons were across age levels.

Multiple Stepwise Regression Analysis

A multiple stepwise regression analysis was computed for each age level for the variable, total arithmetic score. Results are reported in Table 6.

At the seven-year level, seriation accounted for 52.75% of the variance, while inclusion of classes accounted for an additional 12.64%. Both tasks require conservation ability and operational thought. Three other Piagetian tasks, conservation of length, number conservation and concepts of left and right, accounted for 8.90% for a total of 74.31% of the variance. WISC



TABLE 6  
MULTIPLE STEPWISE REGRESSION ANALYSIS

C.A. Group	Variable	Percent of Variance	Cumulative Variance
VII Years	Seriation	52.75	52.75
	Inclusion of Classes	12.64	65.39
	Total Motor Score	3.50	68.89
	WISC: Information	3.49	72.38
	Length Conservation	3.24	75.62
	WISC: Object Assembly	3.21	78.83
	Number Conservation	3.06	81.89
	Concepts of Left and Right	2.62	84.51
IX Years	WISC: Vocabulary	61.16	61.16
	Measurement of Length	15.19	76.35
	WISC: Digit Span	3.57	73.92
	WISC: Coding	3.38	83.30
	DTVP: Position in Span	3.33	86.63
	WISC: Block Design	1.86	88.49
XI Years	Concepts of Left and Right	34.40	34.40
	WISC: Verbal IQ	31.26	65.66
	WISC: Digit Span	6.32	71.98
	WISC: Arithmetic	5.83	77.81
	DTVP: Figure Ground	2.77	80.58
	Oseretzky: Speed and Precision	2.76	83.34
	DTVP: Position in Space	1.51	84.85

Information and Object Assembly, DTVP Spatial Relations, and total motor score accounted for 12.41% for a cumulative percentage of 86.85%.

At the nine-year level WISC Vocabulary accounted for 61.16% of the variance and measurement, a Piagetian task, for 15.19%; that is, two tasks accounted for 76.35% of the variance. WISC Digit Span, Coding and Block Design and DTVP Position in Space increased the cumulative percentage to 88.49%.

At the eleven-year level concept of opposition of left and right, a Piagetian task, and WISC verbal IQ accounted for 34.40 and 31.26% of the variance respectively. WISC Digit Span and Arithmetic accounted for 12.15%, speed and precision for 2.76% and DTVP Figure Ground and Position in Space for 4.28% for a total cumulative percentage of 84.85%.

Conclusions

It would appear that verbal intelligence as measured by psychometric tests is an essential factor in acquiring an understanding of arithmetic and the requirement holds across the age range from seven to eleven years. Achievers emerged as children functioning at least in the average range on all tasks administered. There was marked superiority in verbal skills and average ability in nonverbal skills. Most accelerated were WISC verbal IQ, Vocabulary and Information, all “good” measures of general intelligence (Cohen, 1959). Ability in verbal intelligence tended to separate both achievers and underachievers from nonachievers. The finding is consistent with conclusions by Koppitz (1971).

Logico-mathematical concepts, as measured by Piagetian tasks, appeared



to play an essential role in learning arithmetic but mainly at seven years and to a lesser extent, at nine years. Operational thought in seriation emerged as the most important ability at seven years, followed by inclusion of classes and length and number conservation. Seriation and/or measurement were also involved to a significant degree at nine and eleven years. The pattern appears to be one of slow development towards operational thought, but continuing arithmetic disability. In the early years, operational thought in logico-mathematical concepts appears to be essential for initial success.

Bannatyne (1968) described a spatial organization pattern which includes scores from WISC Picture Completion, Block Design and Object Assembly. The subtests, both alone and in combination, did not differentiate between arithmetic achievement levels in the present study. Bannatyne (1968) also described a sequencing pattern which included WISC Digit Span, Coding and Picture Arrangement subtests. Neither Coding nor Picture Arrangement emerged as significant in the present study; however, WISC Digit Span and Arithmetic emerged as significant both individually and in combination. According to Cohen (1959), at age 13½ Digit Span and Arithmetic combine to measure a factor reflecting "freedom from distractibility." Both subtests differentiated between achievers and nonachievers at seven and nine years, while Digit Span also differentiated between the two levels at age eleven. At eleven years Digit Span and Arithmetic ranked third and fourth respectively in percent of variance contributed in the multiple stepwise regression analysis.

Koppitz (1971) reported a high incidence of hyperactive-distractible behaviour among arithmetic disabled children. The most frequent comment made by teachers in the present study concerned the prevalence of hyperactivity, particularly among the nonachievers. Although Cohen (1959) reported the factor for older children, the factor may be appearing as early as seven years for disabled children.

Two areas of visual perception, Figure Ground and Form Constancy, emerged as significant but only at seven and nine years. The finding is consistent with conclusions by Sabatino and Hayden (1970) concerning the shift in emphasis away from perceptual strategies after nine years.

Although the average number of "motor errors" increased as the degree of arithmetic disability increased, differences were not significant; however, total motor score ranked third in the multiple stepwise regression analysis for the seven-year-olds, after seriation and inclusion of classes. Ayres (1965) reported definite ties between number concepts and space relations, form perception, eye-motor skills and body visualization, with particular identification with perceptual deficits of the two sides of the body. While form perception was related to arithmetic achievement at ages seven and nine in the present study, Eye-Motor Coordination on the DTVP and motor ability were not related. The Oseretzky Test of Motor Ability does not appear, however, to assess many of the areas of motor development described by Ayres. From casual observation many children were exhibiting disabilities described by Ayres. A more relevant test might have been the Purdue Perceptual Motor Survey (1970); however, this test is not standardized.

Only one representational space task emerged as significant. At eleven years concept of opposition of left and right ranked first in percent of variance contributed. The test is concerned with the notion of opposition of left and

right on one's own body, the body of a person opposite and the conceptual notion that left and right are not absolutes.

Four questions were investigated in the present study. Questions 1 and 2 can be answered affirmatively. Elementary school children classified as achievers, underachievers and nonachievers were characterized by different patterns of development. Achievers functioned at least in the average range in all areas assessed. Underachievers differed from achievers on verbal tasks and in stage of development on logico-mathematical tasks. Nonachievers differed from achievers in almost every area assessed but the greatest differences were found for verbal skills, logico-mathematical concepts, visual perception skills in Form Constancy and Figure Ground, and concepts of left and right. Nonachievers also differed from underachievers in verbal skills but were similar in most other areas.

The largest number of statistically significant differences were found at the seven- and nine-year levels. Children at the eleven-year level tended to be operational in spatial and logico-mathematical concepts but continued to have difficulties in learning arithmetic. Verbal intelligence as measured by the WISC seemed to become more closely related to arithmetic achievement as chronological age increased, possibly because changing arithmetic content places greater demands on verbal intelligence, or children may shift from perceptual to conceptual problem-solving strategies as they grow older.

Question 3 (Are one or more of the measures of spatial development related to arithmetic achievement?) can also be answered affirmatively. Perception of form and figure ground were related but at ages seven and nine only. One aspect of representational space, concept of opposition of left and right, was related but at age eleven only. Motor ability as assessed in this study was not related, except in one instance; total motor score ranked third in percent of variance contributed at the seven-year level.

Prior to seven or eight years, according to Piaget, children are under the strong influence of perceptual information, even to the point of making false judgments when data seem conflicting. At approximately seven to eight years, children go beyond the realms of perception and make true judgments in spite of confusing perceptual information. According to Sabatino and Hayden (1970), children move away from perceptual problem-solving strategies to conceptual problem-solving strategies sometime after age nine.

In the present study visual perception, and to a lesser extent, motor ability were related to arithmetic achievement at ages seven and nine but in specific areas, perception of form and discrimination of figure ground. One might postulate tentatively that the two are aspects of perceptual space.

Representational or cognitive space apparently was either intact or not prerequisite for arithmetic achievement. Only at age eleven did representational space appear as a significant factor, in understanding the concept of opposition of left and right. At this point, understanding of left and right becomes a complex conceptual notion, that is, operational functioning on this task involves understanding that left and right are relative rather than absolute concepts, since an object can be left or right at the same time depending upon the number and position of elements in the group.

Spatial organization as measured by the WISC was not related to



arithmetic achievement. Presumably, spatial organization as involved in these tasks is primarily a cognitive skill.

A tentative hypothesis is that arithmetic does involve a spatial component, but until age nine the spatial factors involved are specific aspects of perceptual space. Cognitive space does not appear to be related to arithmetic achievement or content in the early years.

Question 4 asked, "Which of the four areas of the natural number system studied are most severely affected at each of the three chronological age levels?" Underachievers at the seven-year level had most difficulty with place value, while nine- and eleven-year-old underachievers had greatest difficulty with computational skills. Nonachievers had more difficulty than both achievers and underachievers in every area assessed, and the differences were statistically significant.

In summary, the psychoeducational approach to educating exceptional children has been criticized most severely for the naive manner in which assessment and programming have been carried out. The criticisms are relevant in many instances. Matching teaching programs and areas of deficit may be overly simplistic until more is known of the demands made by varied academic content at different age and ability levels.

The present study was concerned with the nature of logico-mathematical and spatial deficits manifested by children in various age ranges and levels of competency in arithmetic achievement. The study was limited by the number of children included in each of the nine groups. This limitation should be rectified by additional research with sufficient numbers of subjects to allow for factor analysis of the areas which appear to be related to arithmetic achievement. Only when we have more exact information concerning manifestations of arithmetic disability and more complete data on the demands placed on intellectual functioning by various forms of academic content will we be in a position to truly evaluate the effectiveness of prescriptive teaching.

#### References

- Almy, M., Chittenden, E., & Miller, P. *Young children's thinking: Studies of some aspects of Piaget's theory*. New York: Teachers College Press, 1966.
- Ayres, A. J. Patterns of perceptual-motor dysfunction in children: A factor analytic study. *Perceptual and Motor Skills*, 1965, 20, 335-368.
- Ayres, A. J. Sensorimotor foundations of academic ability. In W. M. Cruickshank & D. P. Hallahan (Eds.), *Perceptual and learning disabilities in children* (Vol. 2). New York: Syracuse University Press, 1975.
- Bannatyne, A. Diagnosing learning disabilities and writing remedial prescriptions. *Journal of Learning Disabilities*, 1968, 1, 28-34.
- Beard, R. M. The order of concept development—Studies in two fields. *Educational Review*, 1963, 15, 105-117.
- Benton, A. L. *Right-left discrimination and finger localization*. New York: Holber Harper, 1959.
- Blackall, B. Abilities related to mathematical achievement in grade III pupils. Unpublished M.Ed. thesis, University of Alberta, 1972.
- Brown, B. G. Tests of development in children's understanding of laws of natural numbers. Unpublished M.Ed. thesis, University of Manchester, 1969.
- Cohen, J. The factorial structure of the WISC at ages 7½, 10½, and 13½. *Journal of Consulting Psychology*, 1959, 23, 285-299.



- Cohn, R. Developmental dyscalculia. In H. Bakin (Ed.), *Developmental disorders of motility and language: Pediatric clinics of North America*, 1968, 15, 651-668.
- Dodwell, P. C. Children's understanding of number concepts: Characteristics of an individual and of a group test. *Canadian Journal of Psychology*, 1961, 15, 29-36.
- Dodwell, P. C. Relations between the understanding of the logic of classes and of cardinal number in children. *Canadian Journal of Psychology*, 1962, 16, 152-160.
- Dodwell, P. C. Children's perceptions and their understanding of geometrical ideas. In M. Roszkopf et al. (Eds.), *Piagetian cognitive-development research*. National Council of Teachers of Mathematics, 1971.
- Frostig, M. *Marianne Frostig Developmental Test of Visual Perception*. Palo Alto, Calif.: Consulting Psychologists Press, 1966.
- Goldschmid, M. L. Different types of conservation and nonconservation and their relation to age, sex, IQ, MA, and vocabulary. *Child Development*, 1968, 38, 1230-1246.
- Goldschmid, M. L., & Bentler, P. M. The dimensions and measurements of conservation. *Child Development*, 1968, 39, 787-796.
- Glasser, A. J., & Zimmerman, I. L. *Clinical interpretation of the Wechsler Intelligence Scale for Children*. New York: Grune and Stratton, 1967.
- Grewel, F. Acalculia. Lecture delivered at the Institute of Neurology, London Queen Square. *Brain*, 1951, 75, 397-407.
- Grewel, F. Acalculie. *Ned Fijdschr. Psychol.*, 1950, 5, 167-219. English summary, 217-218.
- Hood, M. B. An experimental study of Piaget's theory of the development of number in children. *British Journal of Psychology*, 1962, 53, 273-286.
- Hyde, D. M. G. *Piaget and conceptual development: A cross-cultural study of number and quantity*. New York: Rinehart and Winston, 1970.
- Johnson, D. J., & Myklebust, H. R. *Learning disabilities: Educational principles and practices*. New York: Grune and Stratton, 1967.
- Kaliski, L. Arithmetic and the brain-injured child. *Bulletin of the Orton Society*, 1961, 11, 458-466.
- Kinsbourne, M., & Warrington, E. K. The developmental Gerstmann syndrome. *Archives of Neurology*, 1963, 8, 490-507.
- Kohlberg, L., & DeVries, R. Relations between Piaget and psychometric assessments of intelligence. In C. Lavatelli (Ed.), *The natural curriculum of the child*. Urbana: University of Illinois Press, 1974.
- Koppitz, E. M. *Children with learning disabilities: A five-year follow-up study*. New York: Grune and Stratton, 1971.
- Laurendeau, M., & Pinard A. *The development of the concept of space in the child*. New York: International Universities Press, 1970.
- Lovell, K. *The growth of basic mathematical and scientific concepts in children*. London: University of London Press, 1961.
- Lovell, K. Concepts in mathematics. In J. Klausmeier & W. Harris (Eds.), *Analysis of concept learning*. New York: Academic Press, 1966.
- Lunzer, E. A., et al. SSRC Research Project: The development of systematic thinking. Interim Report Covering 1.9.1970 to 31.8.1971.
- McFarlane-Smith, I. *Spatial ability: Its educational and social significance*. London: University of London Press, 1964.
- Mycok, R. W. The relationship between spatial ability and performance on some Piagetian tests. Unpublished M.Ed. thesis, University of Manchester, 1969.
- Overholt, E. D. An investigation of elementary students' achievement in arithmetic as related to the attainment of the conservation of substance. Unpublished doctoral dissertation, University of Kansas, 1964.
- Piaget, J. *The child's conception of number*. New York: Norton and Co., 1941.
- Piaget, J., Inhelder, B., & Szeminska, A. *The child's conception of geometry*. London: Routledge and Kegan Paul, 1960.

- Piaget, J., & Inhelder, B. *The child's conception of space*. London: Routledge and Kegan Paul, 1967.
- Reimer, A. J. A study of first grade mathematics and conservation. Unpublished M.Ed. thesis, University of Alberta, 1968.
- Riggs, F. Verbal and nonverbal conservation and mathematics achievement. Unpublished M.Ed. thesis, University of Alberta, 1970.
- Roach, E. G., & Kephart, N. C. *The Purdue Perceptual-Motor Survey*. Columbus, Ohio, 1966.
- Sabatino, D. A., & Hayden, D. L. Psycho-educational study of selected behavioral variables with children failing the elementary grades. *The Journal of Experimental Education*, 1970, 38, 40-57.
- Stott, D. H. Second revision of the Gollnitz revision of the Oseretzky Test of Motor Ability. *Developmental Medicine and Child Neurology*, 1966, 8, 523-553.
- Tukey, J. W. Comparing individual means in the analysis of variance. *Biometrics*, 1949, 5, 99-114.
- Vernon, P.E. *The structure of human abilities*. London: Methuen, 1950.
- Werner, H., & Carrison, D. Measurement and development of the finger scheme in mentally retarded children: Relation of arithmetic achievement to performance on the finger scheme test. *Journal of Educational Psychology*, 1963, 33, 252-264.
- Wechsler, D. *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation, 1949.
- Wheatley, G. H. Conservation, cardination, and counting as factors in mathematics achievement. In J. Athey & D. O. Rubadeau (Eds.), *Educational implications of Piaget's theory*. Toronto: Ginn-Blaisdell, 1970.

RICHARD B. COHEN

and

ROBERT H. BRADLEY

*University of Arkansas at Little Rock*

## Teaching Superordinate Concepts with Simulation Games

*The effectiveness of simulation games as a method for teaching superordinate concepts was examined in this study. The simulation involved instruction in map skills and was structured in keeping with Case's (1975) discussion of how to integrate developmental and hierarchical aspects of learning in instructional design. One hundred and eighty-three fifth- and sixth-grade students participated in the week-long experiment. The simulation group demonstrated superior performance on map skills and concepts on the delayed posttest, but not the immediate posttest. Members of this group also showed significant improvement in performance from posttest to delayed posttest. Results were interpreted as supporting the hypothesis that simulations can be a useful way of presenting superordinate concepts so as to facilitate learning of lower level skills. Results also appear to corroborate Case's analysis of effective instructional designs. (Dr. Cohen is Coordinator of Laboratory Experiences and Assistant Professor in the Department of Elementary and Early Childhood Education, and Dr. Bradley is Associate Professor of Educational Foundations and Research Associate at the Center for Child Development and Education, at the University of Arkansas at Little Rock.)*

One of the most significant contributions to the analysis of the learning task has been the development of theories about the hierarchical structure of knowledge (Ausubel, 1963; Gagne, 1970; Taba, 1967; and Bloom, 1956). Conceiving knowledge as hierarchical has enabled educators to analyze tasks according to level of learning and to prescribe a sequence of educational activities which provides the necessary prerequisites for acquiring information at any given level.

Several theorists have argued that when individuals know a superordinate concept, it is easier for them to learn subordinate concepts and facts. Ausubel (1963) advocates the use of advance organizers presented at a higher level of abstraction as a means of facilitating learning at lower levels. However,



advance organizers apparently do not always provide the kind of subsumer that facilitates the learning of subconcepts and specific facts (Barnes & Clawson, 1975). This outcome may partially reflect a limited consideration of the developmental aspects of learning in Ausubel's theory. Case (1975) contends that most learning hierarchies do not provide an adequate analysis of learning tasks, particularly for children. He presents evidence that any analysis of a learning task must take into consideration that young children are more likely than adults to respond intuitively in problem-solving situations. They rely on innate response tendencies which prevent them from adequately evaluating phenomena. To accomplish certain tasks children must learn to respond logically rather than intuitively. Case (1975) feels that to be successful, instruction for young children should deal directly with

both the correct response and the incorrect response. . . . In terms of a hierarchical learning theory such as Gagne's, one would say that the child who participates in the successful training acquires a superordinate skill (or "concept," "plan," or "structure") for attacking the problem which allows him to integrate and discriminate two competing lower-order procedures for attacking it. After training he can succeed consistently, not merely because the correct response has been strengthened, but because response competition has been eliminated. Similarly he can transfer what he has learned, but because any new situation that elicits the incorrect response will also elicit the superordinate or mediating structure into which it has become embedded. Finally, he can retain what he has learned, since he has ample opportunity to apply it (at the supper table, for example, or at play) and little opportunity to confuse it with any other response. (p. 74)

Instructional strategies which consider the developmental nature of learning and which are aimed at teaching superordinate concepts offer a potentially useful way to facilitate the learning of subordinate concepts and facts. One method which can provide a framework for acquiring numerous subskills is simulation games. Simulation games may be defined as an activity in which participants who are attempting to achieve specific goals interact within an artificially produced environment which recreates some aspect of social reality. For the purpose of reality, players assume the roles of individuals or groups who exist in the particular social system being simulated. Bruner has argued that simulation games with their emphasis on participation, informed guessing, hypothesis making and conjectural procedures "go a long way toward getting children involved in understanding language, social organization, and the rest; they also introduce . . . the idea of a theory of these phenomena" (Bruner, 1966, p. 92).

Previous research has demonstrated that simulation games are effective in teaching factual knowledge and concepts. However, Coleman, Livingston, Fennessey, Edwards, and Kidder (1973) explain that simulations tend to be weak in that many participants do not generalize from the particular experiences provided in the game to a general principle applicable in other circumstances. They concluded that learning in school might be made considerably more effective by the "appropriate mix of experiential and information-processing modes of learning" (p. 6). Therefore simulations that incorporate both the symbols helpful for generalization and the actions useful for application to particulars may be the most effective.

A model which may be particularly useful for teaching superordinate concepts with simulation is Taba's instructional model. Taba (1969) identifies

three cognitive tasks involved in the thinking process. The first, concept formation, involves organizing information into a system of classes or groups. In this phase it is sufficient that the facts be introduced and that the learner become somewhat familiar with them. The second cognitive task is interpretation of the data. To do this the learner must relate various kinds of information given and determine cause and effect relationships. Time must be taken to develop concepts, to develop strategies for coordinating lower level concepts and facts, and to make decisions among various alternatives based on the information available. The third cognitive task is problem solving. The learner must be able to apply what he knows, facts and generalizations, to the solution of a problem. This final phase requires additional coordination of facts and concepts, and the development of new strategies for evaluating them.

The purpose of the present study is to examine the effectiveness of simulation as a means of teaching superordinate concepts necessary for the acquisition of subordinate skills. If Case (1975) is correct, then students should demonstrate a knowledge of subordinate concepts when tested at the completion of training on superordinate skills. The encounter with subskills on the test should also constitute a kind of learning trial for the learner. Thus, when tested a second time, student performance should be better. Improved performance would not be expected from students who learned the subskills without benefit of learning the superordinate concepts initially. Previous research on retention indicates that such students are in fact likely to forget some of what they have learned.

### *Method*

#### *Subjects*

Fifth- and sixth-grade classes from two Kansas City, Kansas elementary schools were utilized for this investigation. The eight classes participating in this experiment were randomly assigned to experimental and control groups. The total number of students in the sample was 183, including 91 in the experimental group and 92 in the control group.

#### *Treatment*

The investigator met with the experimental and control teachers one week prior to beginning the treatment. At this time both groups received the same instructions on the purpose of the study and testing procedures. Experimental and control group meetings were conducted separately.

Teachers in the control group were instructed to teach a map skills lesson which included both facts and concepts in the way they usually teach. Their instructional approaches involved teaching the facts and concepts directly rather than teaching superordinate concepts. The lesson used for the sixth grade was from the textbook *Exploring Regions of the Eastern Hemisphere* and for the fifth grade the lesson was from *Exploring Regions of the Western Hemisphere*, both published by Follett. The lesson was to be taught in four consecutive sixty-minute time periods. Time was allowed after the presentation for discussion and questions.

In the preparation meeting with the experimental group teachers, the game *Phantom Submarine* was played since it has been found that teachers



do a better job of teaching a simulation if they have first played it themselves. At this time the teachers became familiar with the game, discovering the rationale, techniques, and procedures.

The systematic approach of Bruner's and Taba's theories for learning and the approach used with the simulation game *Phantom Submarine* are very similar. It provides opportunities for the teacher to expand pupil participation through the use of peer interaction, role-playing, and decision-making. The primary emphasis is on the need for active rather than passive participation in order to have a meaningful learning experience. The active participation leads to intrinsic motivation that Bruner and Taba mentioned as an important factor in helping children develop greater interest in a subject, which, in turn, leads to longer episodes for learning.

After a lengthy discussion about the game, the timetables were established as follows: (a) On the first day the students in the simulation treatment used the sixty-minute period allowed to familiarize themselves with the game and divide into 5-person groups. Each member of the group had a data card telling some physical characteristic of the floor of the Atlantic Ocean. The simulation began by the group compiling factual knowledge about the floor of the ocean and its currents. This was accomplished by a member of the group showing and reading his card to the rest of the group. As each member read his card, the other members recorded this data on their outline maps. (b) The second day, students were presented a problem which required them to make a decision about where some good places might be to try to bring the submarine up from the bottom of the ocean. They were instructed to use the information included on the maps and consider such factors as water depth, mountain ridges, etc. In considering alternative solutions, certain cause and effect relationships for physical features were discussed. Solutions were discarded or approved on the basis of the relationship among these factors. (c) On the third day, students were presented a final problem: What should be done with the submarine? Students were given roles to play in the context of a meeting at the United Nations. The roles involved a consideration of various economic, political, and ecological concerns. Several general questions were posed to the students in the form: What would happen if? To deal with the questions, students had to apply knowledge of the maps and the generalizations they had formed from the previous two tasks. (d) The fourth day was allowed for completing any unfinished tasks and for discussion of all decisions made by the groups.

### *Instruments*

On the fifth day of the experiment, both the experimental and control groups were administered the posttest, the map portion of the Iowa Test of Basic Skills to measure map skills. Two weeks after the posttest a delayed posttest was given to both groups to measure retention.

### *Results*

Means and standard deviations for the Iowa Test of Basic Skills are presented in Table 1. One-way analyses of variance were used to examine differences in the mean performances of experimental and control groups on the Iowa test. No significant difference between groups was observed on the



TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR IOWA TEST OF BASIC SKILLS

	Posttest	Delayed Posttest
Treatment ( <i>n</i> = 91)		
Mean	16.82	17.58
SD	4.73	5.23
Control ( <i>n</i> = 92)		
Mean	16.19	15.59
SD	5.42	6.23

TABLE 2  
SUMMARY TABLE OF ANALYSIS OF VARIANCE FOR  
IOWA TESTS OF BASIC SKILLS POSTTEST

Source of Variation	Sum of Squares	df	Mean Square	F
Between	18.0726	1	18.0726	.8279
Within	3950.9384	181	21.8284	
Total	3969.	182		

TABLE 3  
SUMMARY TABLE OF ANALYSIS OF VARIANCE FOR  
IOWA TESTS OF BASIC SKILLS OF DELAYED POSTTEST

Source of Variation	Sum of Squares	df	Mean Square	F
Between	182.17	1	182.17	5.512*
Within	5982.43	181	33.05	
Total	6164.60	182		

\* significant at the .05 level.

posttest (See Table 2). However, a significant difference was observed between experimental and control groups on the delayed posttest (See Table 3). Furthermore, the experimental group tended to improve from the posttest to the delayed posttest while the control group tended to decline — albeit neither trend was statistically significant.

*Discussion*

Results from the study indicate that the simulation game was an effective way of presenting superordinate concepts so as to facilitate learning of lower level map skills. Performance on the posttest and delayed posttest indicated that learners in the simulation group not only learned as much as learners utilizing more traditional information-processing approaches, but they continued to improve after completion of the training. This subsequent

improvement resulted in superior attainment on the delayed posttest. The findings in this study, which showed that learning superordinate concepts facilitates the learning of subordinate facts and skills, are similar to findings reported by Case (1975) when he trained children to do maze tasks. After training on superordinate concepts, these children tended to improve upon additional encounters with the maze tasks. Children whose training involved only being rewarded for correct approximations to solving the maze task showed no significant advance once the training was terminated. Thus, the present findings tend to corroborate Case's analysis of effective instructional strategies and expand his findings to the learning of several related subordinate concepts.

Educators have not typically espoused the use of simulation games as a means of teaching facts and basic concepts. It seems, however that simulations which involve some aspects of information-processing strategies can be effective in facilitating these types of learning. These types of simulations may serve as useful advance organizers. While utilizing symbols, they are not dependent on written materials. Like other types of experiential learning they involve actions, actions which appear helpful for applying concepts to particulars. Subordinate concepts will not only become "meaningful" in relation to the subordinate concepts presented, but also in relation to the affective response emitted during the actions. As part of a total instructional strategy, simulations offer teachers a useful addition to a repertoire of instructional designs. They provide a change of pace and have been shown to improve attitude toward learning (DeKock, 1969; Cohen, 1969).

Research needs to be aimed at the long-term effects of simulations in facilitating subordinate facts and concepts. In addition, research is needed to examine how simulations might be more effectively integrated into a total instructional approach for a curricular area. Such questions as what types of simulations are needed, how long should they last, how frequently should they be employed and what other types of instructional procedures do they best integrate will need to be considered. Additional research also needs to be done on simulations as advance organizers.

#### References

- Ausubel, D. *The psychology of meaningful verbal learning*. New York: Grune & Stratton, 1963.
- Barnes, B., & Clawson, E. Do advance organizers facilitate learning? Recommendations for further research based on an analysis of 32 studies. *Review of Educational Research*, 1975 45, 637-660.
- Bruner, J. *Toward a theory of instruction*. Cambridge, Massachusetts: University Press, 1966.
- Bloom, B. (Ed.) *Taxonomy of educational objectives: The classification of educational goals. Handbook 1. Cognitive domain*. New York: McKay, 1956.
- Case, R. Gearing the demands of instruction to the developmental capacities of the learner. *Review of Educational Research*, 1975, 45, 59-88.
- Cohen, K. The effects of two simulation games on the opinions and attitudes of selected sixth, seventh, and eighth grade students. *Educational Resources Information Center*, May 1969, ED 031 766.
- Cohen, R. B. "Phantom submarine," A simulation learning situation. University of Kansas, 1974.

- Coleman, J., Livingston, S., Fennessey, G., Edwards, K., & Kidder, S. The Hopkins Games program: Conclusions from seven years of research. *Educational Researcher*, 1973, 2, 3-7.
- DeKock, P. Simulations and changes in racial attitudes. *Social Education*, 1969, 33, 181-183.
- Gagne, R. *The conditions of learning* (2nd ed.). New York; Holt, Rinehart, & Winston, 1970.
- Taba, H. *The Taba Curriculum Development Project in Social Studies*. Menlo Park, California: Addison-Wesley Publishing Company, 1969.



S. C. T. CLARKE

and

STEVE HUNKA

*The University of Alberta*

## Comparative Views on School Discipline

*An Alberta survey of the views of nine different groups on school discipline was conducted during the 1975-1976 school year. Sponsored by the teachers' and trustees' organizations and the Planning and Research branch of the Department of Education, participation reached an overall rate of 77.6% of potential.*

*In line with the results of other similar surveys, considerable satisfaction with the existing state of school discipline was found. However, if there were to be any change, participants generally called for stricter discipline. A difference of interpretation of similar results than that usually associated with those of the Gallup poll is presented, along with findings on traditional versus humanistic views about control of student behaviour, findings about the generation gap between parents and offspring, and dimensions of attitudes toward school discipline. (Dr. Clarke is Professor Emeritus at the University of Alberta, and Dr. Hunka is Coordinator of the Education Research Division in the Faculty of Education at the University of Alberta.)*

Dictionary definitions of discipline include four major elements: education, training or instruction; authority, obedience, subordination, control, direction, or submission; punishment, chastisement or correction; and a resulting state of order or control. School discipline is a narrower concept which generally includes three elements: goals or standards of behaviour, procedures for attaining these goals or standards, and the resulting state. Thus Dettman in the report of an Australian study, defines discipline as follows:

*The discipline of a school is the state or condition of order or good behavior among the students. The term also refers to the procedures by which this state of order is maintained in the school. (1972, p. 7)*

As Smith observes, control of behaviour is essential in society, and as far as schools are concerned, this control is generally recognized in statutes and regulations.

The proper functioning of any social system requires some regulation of its members. . . . For a school system to function properly the conduct of pupils must conform to conditions that are conducive to learning. Principals and teachers are by statutes and board regulations charged with the responsibility for maintaining such order in the school. (1969, p. 292)

Usually, there is a provision in law for school boards to make regulations for managing schools, and frequently four areas of control of student behaviour are specifically mentioned in legislation: expulsion, suspension, corporal punishment, and conduct on school buses. The concept of *in loco parentis*, which means that when a parent sends a child to school the parent is presumed to have delegated to school officials the right to make and enforce reasonable rules and regulations, is one basis for school discipline. Barger (1961) deals with the legal status of students in general, and with corporal punishment in some detail. The conclusions from the extensive literature on this aspect of school discipline are presented in a book by Chamberlin (1971), and in articles by Church (1963), and Clarke and MacKenzie (1970).

There have been many recent surveys of school discipline. Bevan (1973), Chalmers (1972), Hodgkinson (1974) each did studies in a particular Canadian city; Fritz (1973) and MacMillan (1973) in a province; and Lauwerys (1973) in a Canada-wide survey included items related to school discipline. Major findings from these Canadian surveys are presented below. The Australian study by Dettman (1972) previously cited is complete and scholarly. The annual Gallup Poll in the United States as reported in Kappan also includes references to school discipline. Some results from the 1976 poll will later be compared with those from an Alberta study (Clarke, 1977).

A number of studies have applied both scientific knowledge and conventional wisdom to classroom management. Considerable research has centered on the application of behaviour modification as exemplified by Bandura (1969), Becker, Engleman, and Thomas (1971), Glavin (1974), Keller and Inesta (1974), MacMillan (1973), and Sulzer and Mayer (1972). Meacham and Wersen (1974) undertook to integrate behaviourism and humanism. Humanism as applied to education is much broader than classroom management, with Combs (1972), and Zahorik and Brubaker (1972) typical exponents. Kounin (1970) is unique in that his proposals about classroom management are based on research. Chamberlin (1971) has produced a well-received book, as have Jessup and Kiley (1971), the latter applying to the high school level.

A Canada-wide study of the purposes of education was conducted by Lauwerys (1973) for the Canadian Education Association. During the 1972-73 school year 1,680 questionnaires were distributed by this Association across Canada. Lauwerys (1973) reports that of 1531 useable returns there were 236 from students, 257 from educators, 699 from parents with children in school, and 339 members of the general public without children in school. This study found that 21.0% of Canadians checked "lack of discipline" as "the biggest problem with which your schools have to deal." Interestingly, 29.7% of the Alberta respondents checked this. As has been found in other studies, concern about discipline was most pronounced in the general public without children in school, followed by that of parents with children in school, followed by that of educators, and exhibited least by students. Another question asked whether the atmosphere of the school was too restrictive (11.9% checked this response),



too permissive (26.6% checked this), or satisfactory (60.2% checked this). Alberta results showed a smaller proportion checking "satisfactory" (49.7%) and a larger proportion checking "too permissive" (41.3%)

A scholarly survey by Hodgkinson (1974) of the junior secondary schools of Greater Victoria used questionnaires and interviews to gather data. His study found that 90% of administrators, 72% of parents, 64% of teachers, but only 36% of students perceived school discipline as major problem. He pointed out that the widest gap existed between parents and their offspring, and that there was no overall support for a permissive approach. The Lethbridge Public School Board, in a report by Bevan (1973) indicated similar findings. A proposal to eliminate almost all rules and have schools operate on mutual trust and understanding was supported by students but not, in general, by educators, parents, or members of the general public. The report concluded that "there was a continuing desire for organization and structure within the schools." Chalmers (1972) completed a survey entitled *1,000,000 Opinions* for the Greater Montreal Protestant School Board. Students, teachers, and parents rated control measures too strict, about right, or too relaxed. Students rated control measures on seven issues (eg., dress, homework, etc.) as too strict, teachers rated them about right on six of the seven issues, and parents, more than did the other groups, rated control as too relaxed.

From these and other studies, it was clear that in the early 1970's school discipline was a matter of concern. Parents and the general public did not favour any relaxation of control of student behaviour. Teachers, and parents of children in school were, to a considerable extent, satisfied with school discipline as it was, but if there were to be any change, it should be that control should be stricter. However, perhaps the most important finding from all studies was the diversity of views on school discipline. While different publics had, in general, different perceptions about school discipline, the range of views in any one public was great. Majority support for any end state or specific procedure of school discipline was rare.

#### *Research Design, Alberta School Discipline Study*

In the 1975-76 school year, a study of school discipline was sponsored by The Alberta Teachers' Association, The Alberta School Trustees' Association, and Alberta Education, and was funded by the latter. It was also supported by The Alberta Federation of Home and Schools and the Conference of Alberta School Superintendents.

A sample of 100 schools, stratified by size, grade level, and geographic location to be representative of the 1,277 schools in the province was selected, and in fact 100 schools with the required characteristics did participate. Associated administrative units and regional offices were also included. Data were collected by means of a one-sheet (two-page) questionnaire which was identical for each of the participating groups. There were two open-ended questions on the state of school discipline: what *is* discipline like in your school(s)?, and what *should* discipline be like in your school(s)? Connected with the first was a rated type item, on a seven-point scale: 1—altogether too strict, 2—too strict, 3—just about right, perhaps a bit too strict, 4—just about right, 5—just about right, perhaps a bit too lenient, 6—too lenient, 7—altogether too lenient. The rated type item connected with the second was: 1—



much more strict, 2—more strict, 3—perhaps a bit more strict, 4—about as it is now, 5—perhaps a bit more lenient, 6—more lenient, 7—much more lenient. There was another open-ended item asking participants to list one or two causes of school discipline problems. Finally, there were nine suggestions for improving school discipline to be rated on a seven-point scale from 1 (excellent) to 7 (no good, forget it). In total, there were eleven items to be rated, the last nine being as follows:

3. Make schoolwork more interesting and more what kids want to study.
4. After elementary school, let any student of any age who wants to quit school do so.
5. Make sure that students do their schoolwork properly and on time.
6. Give special help to students who cause discipline problems, if necessary putting them in special classes or even schools.
7. Give students more freedom, accept what they do, encourage them to develop self-discipline, and accept some misbehaviour as being part of growing up.
8. Encourage parents to back up the school on what it is doing to get school discipline.
9. Give special help to teachers who are having discipline problems such as special conferences, group sessions, assistance as required.
10. Do whatever is necessary or use whatever measures are necessary (including using the strap) as much as is necessary to keep discipline.
11. Let students help to make and run school rules (guides to conduct).

Participant anonymity was insured by using confidentially marked envelopes. The responses to the open-ended questions were classified by four persons following a carefully prepared key. The ratings of the three largest groups (teachers, parents, students) were factor analyzed to ascertain whether common factors were involved.

### *Results*

The response to the questionnaires was excellent. There were useable returns from 5,228 participants: 100 principals, 80 vice-principals, 1,309 teachers, 1,406 students, 1,715 parents, 48 superintendents, 233 central office personnel, 273 trustees and 64 regional office personnel. Participation ranged from 72% to 100% in the various groups, with an overall rate of 77.6%.

The major findings are summarized below and in greater detail in the *General Report of the Alberta School Discipline Study* (Clarke, 1977). As Table 1 reveals, the responses of the participants were divergent. They spread across the spectrum of possible views on most matters of school discipline. Every group, and all participants together, exhibited this divergence of view. The reaction of the largest single number of participants, a plurality of 41 percent of all participants, was that school discipline as it was in the 1975-76 school year was just about right. Parents especially (47 percent) endorsed this view. Second to satisfaction with school discipline as it was, the next largest number of participants (32 percent) expressed the view that school discipline was too lenient. Consistently, in responding to what school discipline should be like, 34 percent called for it to be stricter.

A number of suggestions for improving school discipline were rated by the participants. The one which exhibited least divergence and most unanimity of view was opposition to the suggestion that after elementary school, any

TABLE 1  
RATING OF PRESENT STATE OF SCHOOL DISCIPLINE

Participants	Rating							Totals
	too strict		about right			too lenient		
	1	2	3	4	5	6	7	
Teachers	3 0.3%	7 0.6%	38 3.2%	444 37.2%	461 38.6%	201 16.8%	39 3.3%	1193
Students	33 2.5%	68 5.0%	327 24.3%	551 41.0%	276 20.5%	71 5.3%	18 1.3%	1344
Parents	3 0.2%	19 1.2%	86 5.5%	728 46.8%	470 30.2%	203 13.0%	47 3.0%	1556
Principals	0 0.0%	1 1.0%	3 3.1%	54 55.7%	37 38.1%	2 2.1%	0 0.0%	97
Vice-Principals	0 0.0%	1 1.3%	2 2.6%	36 47.4%	30 39.5%	4 5.3%	1 1.3%	76
Superintendents	0 0.0%	2 4.3%	6 12.8%	16 34.0%	22 46.8%	1 2.1%	0 0.0%	47
Central Office	0 0.0%	1 0.5%	18 8.2%	70 31.8%	94 42.7%	33 15.0%	4 1.8%	220
Trustees	0 0.0%	0 0.0%	5 1.9%	63 24.2%	125 48.1%	63 24.2%	4 1.5%	260
Regional Office	0 0.0%	4 7.5%	2 3.8%	17 32.0%	21 39.6%	9 17.0%	0 0.0%	53
All Participants	39 0.8%	103 2.1%	487 10.1%	1979 40.9%	1536 31.7%	587 12.1%	113 2.3%	4844

student of any age who wanted to quit school should be permitted to do so. Suggestions that students who were discipline problems, and teachers who were having discipline problems, should be provided with help were rated good to very good. An inducement, “making schoolwork more interesting and more what kids wanted to study” was considered a good idea by students but received lukewarm support from other groups. Another, that students participate in making and enforcing school rules, received little support even from students, and at the elementary school level was deemed inappropriate, especially by parents and teachers.

Following the suggestions for improving school discipline which were to be rated, there was a space for a write-in suggestion. Over 1,000 participants availed themselves of this opportunity, with the most frequent suggestion calling for participation (146 write-ins) followed by standards (123), school conditions such as good extracurricular programs (89), and understanding and accepting students (67). A classification of open-ended responses indicated that the cause of discipline problems most frequently mentioned was students who would not behave properly: they lacked respect for authority and property, exhibited poor attitudes, lacked self-discipline, etc. The second most frequently mentioned cause was lack of discipline in the home. The location of 70 percent of the 9,110 causes listed by the participants was in the school. Another 21 percent of the causes of school discipline problems were located outside the school: in the home, society, etc. The remainder were located in the relationship between the home and the school.



The views of teachers on school discipline were, on the whole, remarkably uniform. Female teachers were more opposed to letting students quit school, and were more in favour of encouraging parents to back up the school, and helping teachers who were having discipline problems than were male teachers. Beyond these differences, there were no other clearcut results. Teachers over 30, as might be expected, held more traditional, conservative, or tough-minded views on school discipline than did teachers 30 years of age and younger. Similarly, mathematics and science teachers held more traditional, conservative, or tough-minded views than did English and social studies teachers. The same kinds of differences were exhibited between classroom and non-classroom teachers, and between classroom teachers and guidance counsellors.

Students who aspired to matriculation, college or university were much more traditional, conservative, or tough-minded in their views on school discipline than were the non-matriculation and non-college bound students. Paradoxically, the first group was also more in favour of humanistic measures of student behaviour control. These findings support the conventional wisdom that it is the nonacademic student who turns to misbehaviour.

There was a marked difference in views on school discipline in parents over 30 years of age in contrast with those 30 years of age or younger. The former were more traditional, conservative, or tough-minded. A similar difference was exhibited between parents of "two-parent" and "one-parent" families.

There were a number of expected group differences in views on school discipline. The views of teachers and trustees were close together on what discipline *was* like and on what it *should be* like. At the opposite pole were students and superintendents. For further illustrations of group differences, the reader is referred to Clarke (1977). Although groups differed one from another in their views on school discipline, group differences were overshadowed by the diversity of view within any one group.

An important finding of the study was that for elementary schools, in the view of participants with respect to school discipline, big was bad and small was good. That is, teachers, parents, and other participants associated with large schools rated discipline as being worse than did the corresponding groups associated with small schools.

The following sections provide somewhat more detail on four other findings: comparisons with Gallup Poll results, traditional vs. humanistic attitudes to student control, the generation gap, and dimensions of attitudes toward school discipline.

### *Comparisons With Gallup Poll Results*

The eighth annual Gallup Poll of the American public's attitudes toward the public schools as reported in the October, 1976 issue of Kappan continues to reinforce the impression that discipline is the number one problem. The report uses these words:

Discipline continues to head the list of major problems when a sample of the nation's adults cite what they perceive as the most important problems of the public schools in their own communities. In fact, discipline has been named most often seven times during the last eight years.



TABLE 2  
COMPARISON OF GALLUP POLL AND ALBERTA FINDINGS

Attitude toward discipline	Response
Discipline is poor (Alberta)	27%
Lack of discipline (Gallup)	22%
Discipline is good (Alberta)	55.9%
Discipline is good (Gallup)	--

Newspaper headlines, which catch the phrase, "Discipline, the schools' number one problem" have probably implanted the idea in the minds of many.

As Table 2 reveals, some findings of the Alberta study were similar to those of the American study, but the latter did not provide an opportunity for participants to rate the present state of school discipline. The comparisons should not be pushed too far because the questions asked were not identical, the samples differed, and the classification of open-ended responses is always a matter of interpretation in borderline cases. However, as can be seen from Table 2 as well as from an examination of the results of the rating type items as presented in Table 1, the Alberta respondents exhibited considerable satisfaction with school discipline. This is a far cry from the impression created by the headlines, "Discipline, the school's number one problem."

#### *Traditional vs. Humanistic Attitudes to Student Control*

Carlson (1964) proposed a theoretical framework which was developed by Willower (1967) and his associates at Penn State. They developed a Pupil Control Ideology instrument which was used by Fritz (1973) in Alberta and MacMillan (1973) in Nova Scotia. This instrument was not used in the Alberta School Discipline study but two items, to be rated on a seven-point scale as suggestions for improving school discipline, were designed to tap the theoretical positions mentioned.

The item designed to tap the traditional position (called custodial by the originators) was as follows: "Do whatever is necessary or use whatever measures are necessary (including using the strap) as much as is necessary to keep discipline." Responses to this item showed that participants were markedly divided. Thus 13.6% rated it 1 (excellent, tops) and at the other extreme, 14.5% rated it 7 (no good, forget it). There was 13.8% who rated it 2 (very good), 14.5% who rated it 6 (little good), and 16.0% who rated it 3 (good) while 9.5% rated it 5 (might be some good). Only 18.1% selected the neutral or middle position. Teachers and trustees favoured this suggestion, while regional office personnel and central office personnel opposed it (followed by students!).

The item designed to tap the humanistic position was as follows: "Give students more freedom, accept what they do, encourage them to develop self-discipline, and accept that some misbehaviour is part of growing up". The response of participants showed divided views with the majority opposed.

The ratings were (see above for descriptions): 1, 7.9%; 2, 10.8%; 3, 16.1%; 4, 19.3%; 5, 11.9%; 6, 17.0%; 7, 17.1%. Students and regional office personnel favoured this suggestion while parents and trustees opposed it.

It is clear from the above results that no consensus was present about traditional measures for student control. In general, the traditional mode of student behaviour control received majority support, but sizeable proportions of participants held strong views both for and against. The humanistic position on student behaviour control measures as expressed in the item used was not supported by a majority. Again there was divided opinion, with sizeable proportions strongly opposed and somewhat smaller proportions strongly in favour.

*The Generation Gap*

Of the 1,406 students and 1,715 parents who participated in the study, 1,338 were matched, parent with offspring. The correlations between student and parent ratings on the eleven rating scale type items are shown in Table 3. These correlations are all low, but they are all positive. The greatest tendency for a high parental rating of an item to be associated with a high student offspring rating of the same item (or low with low) was demonstrated in item 10, traditional measures for control of student behaviour, where  $r = 0.310$ . The least such tendency was exhibited in item 6, provide special help for students who cause discipline problems, where  $r = 0.139$ .

This may be a surprising result for those who think in terms of a generation gap, because it supports the old saying, "like parent, like child." This tendency is present for all eleven items.

Table 3 also demonstrates the generation gap. It shows the mean position of the 1,388 parents compared with the mean position of their 1,388 offspring on the various items. All except one of the differences shown was significant at the 0.05 level of confidence. With respect to item 1, parents rated school

TABLE 3  
COMPARISONS OF PARENT-OFFSPRING MEAN POSITIONS  
ON SCHOOL DISCIPLINE

Number and Paraphrase of Items	Mean Parents	Mean Students	Correlation
<u>Is and Should Be</u>			
1. What discipline is like	4.24	3.84	0.255
2. What should discipline be like	3.01	3.86	0.209
<u>Suggestions to Improve/Discipline</u>			
10. Traditional measures for discipline	3.43	4.43	0.310
4. Let students quit after elementary school	6.07*	6.05 *	0.279
3. Make schoolwork more interesting	3.20	2.66	0.275
7. Humanistic measures for discipline	4.49	3.54	0.223
8. Encourage parents to back up school	1.98	3.44	0.216
11. Let students help run school	3.79	3.57	0.215
9. Special help to teachers	2.37	3.16	0.163
5. Schoolwork done properly and on time	2.10	3.21	0.161
6. Special help to students	2.84	3.56	0.139

\* difference not significant at .05 level of confidence.



discipline as more lenient than did their offspring, and consistently, in item 2 indicated by their rating it should be stricter to a greater extent than did their offspring. In the suggestions for improving school discipline, parents rated item 10, traditional measures for control of student behaviour, much more favourably than did students. The balance of the items under this heading are interpreted in the same way: lower numbers mean a more favourable rating. All of the differences are in the expected directions, and all are statistically significant at the 0.05 level of confidence except one.

These results were seemingly a paradox. There was a positive correlation between parents' views and those of their offspring. Like parent, like child. Yet there was a pronounced generation gap, in the expected directions on the various items. In general, parents favoured stricter control and requiring students to perform. The findings of the study support both positions.

### *Dimensions of Attitudes Toward School Discipline*

The eleven items which participants rated were factor analyzed for three groups: teachers (1,309), students (1,406), and parents (1,715). A principal factor analysis with squared multiple correlations used as communality was performed separately for each group. The results are shown in Table 4.

Factor I, TRADITIONALISM IN TRANSITION, includes these items:

5. Make sure that students do their homework properly and on time.
8. Encourage parents to back up the school on what it is doing to get school discipline.
10. Do whatever is necessary or use whatever measures are necessary (including using the strap) as much as is necessary to keep discipline.
6. Give special help to students who cause disciplinary problems, if necessary putting them in special classes or even schools.

TABLE 4  
VARIMAX FACTOR LOADING MATRIX

Item	Factor	Teachers		Parents		Students	
		I	II	I	II	I	II
5		0.48	-0.11	0.52	0.05	0.56	-0.01
8		0.50	0.02	0.62	0.08	0.55	0.11
10		0.46	-0.22	0.58	-0.03	0.61	-0.09
6		0.41	0.09	0.46	0.27	0.49	0.14
9		0.40	0.25	0.46	0.29	0.45	0.26
3		0.14	0.51	0.10	0.53	0.06	0.44
11		0.20	0.45	0.19	0.53	0.09	0.52
7		0.07	0.59	0.09	0.66	-0.01	0.57
4		0.14	0.02	0.33	0.48	0.22	0.41
1		-0.16	0.41	-0.07	0.38	-0.28	0.42
2		0.19	-0.35	0.26	0.04	0.43	-0.33
Percent of Total Variance		11.6%	10.2%	20.6%	8.3%	16.0%	12.1%



9. Give special help to teachers who are having discipline problems such as special conferences, group sessions, assistance as required.

A common thread running through these items is an emphasis on standards, upholding standards, and on excellence. It reflects the position that the school is responsible for ensuring student learning; traditionalism tempered with compassion. The factor was identified as one of "traditionalism in transition."

Factor II, HUMANISM, includes these items:

3. Make schoolwork more interesting and more what kids want to study.
11. Let students help to make and run school rules (guides to conduct).
7. Give students more freedom, accept what they do, encourage them to develop self-discipline, and accept that some misbehaviour is a part of growing up.
4. After elementary school, let any student of any age who wants to quit school do so.
1. What do you think about discipline in your school?

A common thread running through these items is freedom for self expression, freedom to grow. It is laissez-faire with social compassion; the school is responsible for permitting students to develop. It includes the idea of helping the student to unfold. It calls for freedom to grow. The factor was identified as "humanism." The two factors which appeared account for one-quarter to one-third of the variance exhibited. It is important to note that the factors apply to each of the groups: teachers, parents, and students. The one exception is that item 4 above, for teachers, does not belong in Factor II, Humanism. It may be that teachers who were in favour of students quitting school were, in fact, traditional in their attitudes.

Item 2 was, "What should discipline be like in your school?" This question, for both students and teachers is related to both traditionalism in transition and to humanism. However, for parents, this item appears not to be in those domains. That students' attitudes toward school discipline are in the same domain as are teachers' and parents' is a matter of note. It may be that students can accept the traditionalism in transition position so long as it is fair. Traditionalism has often been equated with custodialism, and traditional attitudes about student control in school linked, as was the traditional-humanistic control orientation, to the attitudes associated with control in prisons and mental hospitals. While this may have held in the past, the present study finds a less restrictive and more compassionate component. Scholarship, excellence, and standards remain, but they are tempered with social compassion.

Humanism may be an emerging or evolving attitude toward student control. In the words of youth, it is "do your own thing," or in the words of Abraham Maslow, "what a man *can* be, he *must* be" (self-actualization). Even students accept the traditional position (excellence) and accept that behaviour must be controlled, either internally or externally. Their one caveat seems to be that control is fair.

### Conclusions

The findings of the the present study indicate that there was considerable satisfaction with the state of school discipline at the time the study was made.

However, there was support for stricter disciplinary procedures and, consistently, little support for less restrictive control measures.

In general, it has been found that the public at large is most critical of school discipline, followed by parents and school personnel. In the Alberta study, teachers and trustees were more strongly in favour of stricter discipline than were parents of students. While there is no doubt that grade level and community expectations affect views on school discipline, the general findings support the view that an overriding factor is the expectations people hold. The temper of the times, or the social context, appears to be the most influential factor determining views on school discipline.

The diversity of view exhibited by participants from any one group was great. This, rather than consensus, was the rule. School disciplines is not a single unitary matter toward which group consensus of viewpoint is exhibited. Rather, there is such diversity of view even within any one group such as teachers, parents, or students that a change in emphasis or procedures would inevitably please some and displease others.

The complexity of school discipline is such that the approach used in any study can predetermine results unless care is taken to provide open-ended responses or sufficient alternatives to cover most of the views participants hold. In terms of the media, contradictory headlines could well be used to highlight the same data.

The strengths of the present study were its province-wide extent, the use of one instrument for all participants, the open-ended items, the carefully chosen representative sample, and the preparation and backing which helped produce the high rate of return of questionnaires. The weaknesses were that it was place-bound and time-bound to Alberta in the mid 1970s, it relied on a questionnaire, open-ended items had to be classified, discipline was not defined for the participants, the groups studied were limited to nine, and the context in which discipline was set (social trends) were not studied.

#### References

- Bandura, A. *Principles of behaviour modification*. New York: Holt, 1969.
- Bargen, P. F. *The legal status of the Canadian public school pupil*. Toronto: Macmillan, 1961.
- Becker, W. C., Engleman, S., & Thomas, D. R. *Teaching: Classroom management and cognitive learning*. Chicago: S.R.A., 1971.
- Bevan, G. H. (Chairman of Committee) *Goals of public education: A report of the Educational Goals Project Committee*. Lethbridge, Alta.: Lethbridge Public School District #51, June, 1973.
- Carlson, R. D. Environmental constraints and organizational consequences: The public school and its clients. In D. Griffiths (Ed.), *Behavioral science and educational administration*. Chicago, Ill.: The University of Chicago Press, 1964, 262-276.
- Chalmers, R. (Project Director) *1,000,000 opinions: A report of the social values survey*. Montreal: The Protestant School Board of Greater Montreal, May, 1972.
- Chamberlin, L. J. *Effective instruction through dynamic discipline*. Columbus, Ohio: Merrill, 1971.
- Church, R. M. The varied effects of punishment on behavior. *Psychological Review*, 1963, 70, 369-402.



- Clarke, A. M., & McKenzie, I. C. Punishment and its behavioral outcomes: Application of research evidence to the modification of behavior. *Australian Journal of Education*, 1970, 14 (1), 30-48.
- Clarke, S. C. T. *General report of the Alberta School Discipline Study*. Edmonton: Planning and Research, Alberta Education, 1977.
- Combs, A. W. Educational accountability: Beyond behavioral objectives. Washington, D. C.: A.S.C.D., 1972.
- Dettman, H. W. *Discipline in secondary schools in Western Australia: Report of the Committee enquiring into discipline in secondary schools in Western Australia under the chairmanship of Mr. H. W. Dettman, Perth, December, 1972*. The Education Department of Western Australia, 1972.
- Fitz, J. O. *Control ideology and the student in Alberta high schools*. Proceedings of the Fifth Western Canada Educational Administrators' Conference. C.S.A. Monograph #4, Council on School Administration, The Alberta Teachers' Association, Barnett House, Edmonton, 1973.
- Glavin, J. P. *Behavioral strategies for classroom management*. Columbus, Ohio: Merrill, 1974.
- Hodgkinson, C. Cross perceptions and cross purposes. The junior secondary years in Greater Victoria. (Unpublished). The University of Victoria, 1974.
- Jessup, M. H., & Kiley, M. A. *Discipline: Positive attitudes for learning*. Englewood Cliffs, N.J.: Prentice Hall, 1971.
- Keller, F. S., & Inesta, E. R. *Behavior modification: Application to education*. New York: Academic Press, 1974.
- Kounin, J. *Discipline and group management in classrooms*. New York: Holt, Rinehart and Winston, 1970.
- Lauwerys, J. *The purposes of education: Results of a C.E.A. survey*. Toronto: Canadian Education Association, 1973.
- MacMillan, M. R. Pupil control ideology and status obeisance of teachers and principals in elementary schools. Unpublished doctoral thesis, Faculty of Education, The University of Alberta, 1973.
- Meacham, M. L., & Wersen, A. E. *Changing classroom behavior* (2nd ed.). New York: Intext Educational Publishers, 1974.
- Smith, B. O. Discipline. *Encyclopaedia of Educational Research*, Fourth Edition. Toronto: Macmillan, 1969.
- Sulzer, B., & Mayer, G. R. *Behavior modification procedures for school personnel*. Hinsdale, Ill.: Dryden, 1972.
- Willower, D. J., Eidell, T. L., & Hoy, W. K., *Penn State studies: The school and pupil control ideology*. University Park, Pa.: The Administrative Committee on Research, The Pennsylvania State University, 1967.
- Zahorik, J. A., & Brubaker, D. L. *Toward more humanistic instruction*. Dubuque, Iowa: Brown, 1972.



## BOOK REVIEWS

APPLYING BEHAVIOUR-ANALYSIS PROCEDURES WITH CHILDREN AND YOUTH. *By B. Sulzer-Azaroff and G. R. Mayer.* Toronto: Holt, Rinehart and Winston, 1977.

Nathan Azrin, in the Foreword to the book, describes the extensive and varied experiences the authors have had that make them "uniquely qualified" to write the book. Their experience reportedly has included public schools from the primary grades through high school, special education classes, residential institutions and parent-child interactions. They have worked with children in regular classrooms as well as with emotionally disturbed, profoundly to mildly retarded, and autistic children.

This wealth of experience is certainly expressed in the variety of examples used in this topically comprehensive, practical and informative book on applied behaviour analysis.

The thirty five units of the text provide a complete, sequential model of applied behaviour-analysis (ABA). The following topics are reviewed:

- a) contemporary legal/ethical issues
- b) selection of goals and objectives
- c) reliability, validity and utility of various data recording methods
- d) the advantages, disadvantages, selection, implementation and evaluation of a variety of intervention strategies to increase, decrease, teach, extend and maintain behaviour
- e) applied behaviour analysis (ABA) research design procedures, and
- f) method for communicating results of ABA programs.

The book is five hundred and two pages in length, plus a brief appendix of behavioural assessment systems, an excellent fifteen-page glossary of technical terms, and a comprehensive, detailed, and well organized subject index.

Performance goals are stated at the beginning of each unit and additional advance organizers are frequently incorporated into the body of the text. These features help to orient the reader, focus his attention on essential issues, and guide his progress through the material. Each unit concludes with a brief summary. Technical terms, defined in the context and in the glossary, are printed in bold type.

An outstanding feature of the book is the use of check lists, tables and flow charts that summarize the material and sequentially list the major points to

facilitate decision-making during selection and application of behavioural techniques. For example, decision-making flow charts are provided for:

- a) practical considerations to make before implementing an ABA program
- b) selecting priorities in establishing goals, and
- c) selecting and implementing recording techniques, procedural strategies and functional analysis designs.

With these features the text will function as a continuing reference for practising applied behaviour-analysts.

A study guide describes field activities designed to provide students with the opportunity to observe, design, implement and evaluate programs. Sets of quizzes and marking keys are available from the publisher upon request.

Throughout the text, the authors emphasize the use of positive, constructive, adaptive, and minimally intrusive methods of behaviour management rather than coercive, suppressive, or merely institutionally expedient procedures. The authors also stress concern for the protection of human rights. They discuss the concepts of voluntary, informed consent and client involvement in treatment selection and evaluation. Emphasis is also placed upon accountability, and ethical and legal responsibility. Several recent U.S. court cases are discussed as are practical methods for avoiding the violation of human rights, litigation, and overly restricted legislation. Although the discussion relates specifically to the United States, the problems described are universal. An awareness of the U.S. experience may give Canadian readers the advantage of repeating some of the successes achieved while avoiding some of the problems.

In the preface to the book the authors describe several ways in which the text may be used:

- a) in conjunction with the study guide and practicums to train professional competencies over one academic year;
- b) to train technical and paraprofessional competencies in a one-semester course;
- c) to train parents, teacher, ward staff and others in a two-semester program; and
- d) to provide inservice training on special topics.

Selected portions of the text are designated for the various programs. The impression that the book provides an answer to most training needs in the use of applied behaviour analysis procedures may be heightened by the use of examples drawn from schools, hospitals, residential facilities, institutions, and clinics working with medical, psychiatric, social, educational and family problems.

The variety of examples cited in the text certainly emphasize the generality of the ABA techniques. However, the variety of examples used may actually impede learning. Optimal generalization of techniques and transfer of principles into practice is achieved when instruction includes examples describing conditions common to the environment in which the learner is expected to perform. The drawing of examples from many different areas of application reduces the number of examples taken from any one area. Student teachers, for example, may have difficulty seeing the educational relevance of "stimulus control" when the examples provided describe a pilot landing in a dense overcast, and a doctor making a diagnosis based on the presence of particular symptoms.



In the foreword to the text, Azrin suggests that the last section of the book on research methods and reporting skills is ideally suited to graduate students and provides an excellent foundation for designing behavioural experiments for theses. In my opinion, the text is suitable as an introduction to applied behaviour-analysis possibly for senior level undergraduates and preferably for graduate students in clinical counselling, community psychology or educational psychology.

Azrin also suggests that the content provides a self-contained cookbook describing in a step-by-step manner what to do. The impression might be given that one has only to read the text to become a competent applied behaviour-analyst. The authors, however, state throughout the text that supervised practice is a prerequisite to competence. They state that "change agents must be adequately trained . . . and they must be regularly supervised . . . training alone is not enough to guarantee that a procedure will be implemented properly" (p. 85).

A supervised practicum is, for three reasons, a necessary complement to the text. First, as discussed, the variety of examples used in the text do not appear optimally to facilitate transfer and generalization. Second, the vast majority of instructional objectives stated at the beginning of each unit are at the knowledge and comprehension level of Bloom's Taxonomy and are not directed to the application of techniques in the solution of practical problems. Thirdly, inadequately trained practitioners may misapply techniques producing failures for themselves and their clients, resulting in the possible rejection of potentially useful methods. Thus, as the authors suggest, the text should be used in conjunction with a supervised practicum.

The text does not have an inspiring beginning. Chapter one begins with an abstract definition of applied behaviour analysis, a discussion of the principles of behaviour, an introduction to technical terminology, a review of the responsible application of behaviour analysis techniques, and a list of the areas in which the methods have been used successfully. Similarly, units two and three describe procedural and ethical considerations. The examples cited describe the hypothetical problems of Flossie, Lucretia, Mr. Grump, and Dexter. Greater reader interest and understanding may have been achieved through the use of an extended description of actual applications of the techniques in the solution of real problems. There are many exciting examples to choose from in the literature that demonstrate definitional, procedural, ethical, and other basic concepts in an interesting and inspiring manner.

The discussion in units two and seven on narrative recording and sequence analysis describe a procedure whereby the sequences of antecedent events, problem behaviours and their consequences are recorded in "running descriptions" over several days of observation. After four hundred events have been recorded "a trend might become manifest." The sheer size of this task is certain to alienate many readers. In most communities there are few people available with the time and expertise necessary to employ this approach in a valid and reliable manner. In addition, the text does not adequately describe how to select, collect, or analyze the data.

The book provides a topically comprehensive, practical approach to the selection, application, and evaluation of applied behaviour management techniques for a variety of problems and situations. The authors emphasize a legally and ethically responsible, positive approach to behaviour management.

The text is clearly written and many of the major points have been



excellently summarized in check lists, tables and flow charts. These devices facilitate organization of newly learned material and will continue to function as practical references during application of the ABA techniques.

Because the authors have drawn their examples from such a broad variety of contexts, students specializing within a particular area, for example Education, will likely require considerable instructional assistance in putting the examples into an educational perspective. Extensive supervised practicum experience will be required to assist the transfer of theory into practice.

The text is suitable for the introductory training of "applied behaviour-analysts" possibly at the upper undergraduate level and preferably at the graduate level during a one-year course as a part of a clinical counselling, community psychology, or educational psychology program.

David Baine  
*Department of Educational Psychology*  
*The University of Alberta*

TEACHING LOW-ACHIEVING CHILDREN READING, SPELLING AND HANDWRITING: DEVELOPING PERCEPTUAL SKILLS WITH THE GRAPHIC SYMBOLS OF LANGUAGE. *By Annabelle M. Markoff.* Springfield, Illinois: Charles C. Thomas 1976.

Although there has been an increasing number of texts which suggest integration of language arts for average and high-achieving students, Annabelle Markoff's book is one of the few recent publications which attempts to delineate an integrated program for low-achieving children. In this regard the title of her book is somewhat reminiscent of Gillingham and Stillman's *Remedial Training for Children with Specific Disability in Reading, Spelling and Penmanship* first published in 1940. The major purpose of Markoff's book, however, is markedly different from the Gillingham-Stillman volume, growing out of her concern with the current emphasis on geometric forms in perceptual training as evidenced in the work of Frostig, Kephart and Kirk. She attempts to provide a rationale for shifting the emphasis in perceptual training away from geometric forms to basic language forms involving sounds, letters, and words. This is a very welcome shift and one which has considerable research support. Her second major purpose is to suggest teaching techniques for integrating perceptual training in the teaching of the subjects themselves.

As evidence of Markoff's concern with providing a rationale for curriculum decisions with low-achieving children, she devotes the first chapter to a discussion of the relationships between reading and spelling. Both are seen as communication processes: in reading, one translates symbols into meaning; in spelling, meaning is translated into symbols. She emphasizes the importance of meaning, and it appears in this initial chapter that she is using spelling as a term synonymous with written composition. However the meaning dimension, although receiving considerable attention in the chapter on reading, is considered only minimally in the chapter on spelling and almost not at all in the one on handwriting. Hence, the last half of the book seems only marginally related to the theoretical background presented in the first chapter. In addition, Markoff does not systematically review the literature when building her theoretical framework, but draws

heavily from research in educational psychology, particularly from E. Gibson's *Principles of Perceptual Learning and Development*. She makes little reference to studies or theories produced by reading educators, and makes many definitive statements about how people read and spell with little attempt to provide research support.

In the chapter on reading, Markoff deals with six areas: perceptual skill, decoding, vocabulary, fluency, comprehension, and motivation. She expresses concern with the decoding-comprehension dichotomy prevalent in the literature, and is particularly concerned about the heavy emphasis on decoding (word analysis) frequently found in programs for low-achieving readers. She correctly suggests, "To teach reading as a decoding process or separate it from . . . the communication process itself, is a short-sighted solution" (p. 94). The remainder of the chapter, however, is somewhat disappointing in terms of both cohesiveness and the relationship of teaching suggestions to theory and research in reading. She relies heavily on the work of Montessorri and Fernald to suggest a multi-sensory approach to teaching letter and word identification. She then turns to a discussion of the language experience story and suggests its use daily with low-achieving students to "supplement, complement and integrate all forms of language instruction." She presents this "integrative vehicle" as the key to developing comprehension, fluency and motivation. There is very little research to demonstrate the superiority of either a multi-sensory or language experience approach with low-achieving children but this does not deter Markoff from making a very strong statement about the usefulness of these techniques. At this point in time judicious use of language experience stories seems to be supported by conventional wisdom, but to claim, as Markoff does, that this has a scientific basis is questionable.

The chapters on spelling and handwriting are, like the chapter on reading, a mixture of ideas gleaned from research and Markoff's own work in special education. She places heavy emphasis on the perceptual aspects of spelling and handwriting, and gives considerable attention to modalities in her discussion. In her suggestions for teaching spelling, she deals largely with isolated words and, with the exception of meaning vocabulary, treats spelling as distinct from the other language arts. In both this and the chapter on handwriting, her emphasis is on the importance of helping children establish distinctive features of letters and sounds. A multi-sensory approach with a heavy motor component is suggested similar to that in the chapter on reading.

Editorially this book has several weaknesses. The order of the chapters noted on the bookjacket does not correspond to that in the text. There are also numerous references within the text to other sections which results in confusion because of this change in order. At several points Markoff refers to an appendix which is nonexistent, and there are a large number of typographical errors. Finally, the literature reviewed in the book is both narrow and somewhat outdated as the latest reference is 1974.

The basic weakness, however, relates to Markoff's attempt to wed theory and practice in teaching low-achieving children. She states in the preface to this book that it started as a series of teaching suggestions for students in special education and regular classroom situations, and that this book was an attempt to organize these ideas into a theoretical framework. She appears to have made a priori curriculum decisions and then to have searched for a theory to support these decisions, rather than making curriculum decisions on the basis of a sound theoretical formulation. This has resulted in

inconsistency among ideas at several points and, in the last two chapters, the addition of another “teaching idea” after the summary to the chapter. Hence, although she provides several useful suggestions, the overall effect has been somewhat less integrated than intended.

Grace Malicky  
*Reading and Language Centre*  
*Faculty of Education*  
*The University of Alberta.*



# THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

## Volume XXIII, 1977 — Table of Contents

- Allen, L. E., and Ross, J. *Improving skill in applying mathematical ideas: A preliminary report on the instructional gaming program at Pelham Middle School in Detroit.* No. 4, December, 257-267.
- Bana, J. P., and Nelson, D. *Some effects of distractions in nonverbal mathematical problems.* No. 4, December, 268-279.
- Bezeau, L. M. *Equality of educational opportunity and inequality of per pupil expenditures.* No. 3, September, 218-225.
- Blowers, E. A. *Prediction of Metropolitan Readiness Test scores.* No. 2, June, 164-168.
- Bond, J. B., Jr. *Change in anxiety level as a factor in test performance.* No. 2, June, 97-102.
- Bourgeois, R., and Nelson, D. *Young children's behavior in solving division problems.* No. 3, September, 178-185.
- Clarke, S. C. T., and Hunka, S. *Comparative views on school discipline.* No. 4, December, 305-316.
- Cohen, R. B., and Bradley, R. H. *Teaching superordinate concepts with simulation games.* No. 4, December, 298-304.
- Cummins, J., and Das, J. P. *Cognitive processing and reading difficulties: A framework for research.* No. 4, December, 245-256.
- Harper, F. B. W. *Practice teaching performance and resultant achievement motivation.* No. 2, June, 104-108.
- Haughey, M., Holdaway, E. A., and Small, J. M. *Consultative practices of elementary school teachers: Some research findings.* No. 1, March, 85-96.
- Hodapp, T. V. *Children's ability to learn problem-solving strategies from television.* No. 3, September, 171-177.
- Hodysh, H. W. *A critical examination of Dewey's method of inquiry.* No. 3, September, 213-217.
- Hundert, J., and Bucher, B. *Beginning readers discriminate words by their simplest cue.* No. 3, September, 186-194.
- Jansson, L. C. *Conditional reasoning in adolescents.* No. 2, June, 118-127.
- Janzen, H. L., and Hallworth, H. J. *Linguistic predictors of properties of set.* No. 1, March, 11-21.
- Kirman, J. *The use of infra-red false color satellite images by grades 3, 4, and 5 pupils and teachers.* No. 1, March, 52-64.
- Konrad, A. G. *College trustees view their selection.* No. 2, June, 138-150.
- Marcus, R. F., and Jenny, B. *A naturalistic study of reciprocity in the helping behavior of young children.* No. 3, September, 195-206.
- Martin, J., and Auerbach, S. *Differential outcome effects in teacher training research as a result of functional versus structural behavioral recording.* No. 2, June, 128-137.
- McCann, S. J. H., and Fisher, R. J. *The interactive relationships of teacher directiveness and student authoritarianism and dogmatism to grades and satisfaction.* No. 1, March, 71-84.



McLeish, J. *Knowledge and revolution: The student movement in Tsarist society (1856-1881)*. No. 2, June, 151-163.

Nelson, L. D., and Kieren, T. E. *Children's behavior in solving spatial problems*. No. 1, March, 22-30.

Noll, R. L., Willower, D. J., and Barnett, J. J. *Teacher self-actualization and pupil control ideology-behavior consistency*. No. 1, March, 65-70.

Phillips, S. *The contributions of L. S. Vygotsky to cognitive psychology*. No. 1, March, 31-42.

Powell, J. C. *The developmental sequence of cognition as revealed by wrong answers*. No. 1, March, 43-51.

Powell, J. C., Cottrell, D. J., and Lever, M. *Schools I would like to see: An opinion survey instrument with interesting possibilities*. No. 3, September, 226-241.

Ross, C. J., and Stewart, L. D. *Graduates' assessment of industrial arts teacher education programs*. No. 2, June, 109-117.

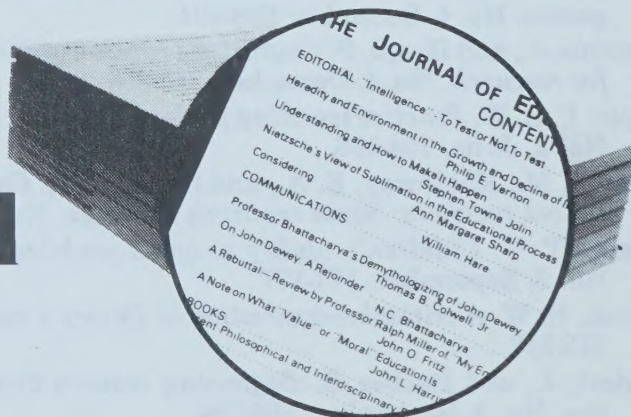
Tyler, J., and Hardy, R. C. *Effects of age and sex on the acquisition of form discrimination*. No. 1, March, 1-10.

Whyte, L. *Logico-mathematical and spatial development in children underachieving in arithmetic*. No. 4, December, 280-297.

Williams, R. J., and Algozzine, B. *Differential attitudes toward mainstreaming: An investigation*. No. 3, September, 207-212.

# The Journal of Educational Thought

... provides an outlet for the discussion of educational ideas. In general, it complements empirical research by speculative, critical thought about public education. The Journal draws specifically upon philosophical, social and historical insights, but is receptive to other approaches both theoretical and practical - including both those of teachers and administrators. It is intended to serve a broad readership which includes academicians, teachers, administrators, and the general public.



## Subscription Rate: Back issues (prior to 1976):

\$ 9.00	1 year	Single issue	\$1.00 per copy
\$17.00	2 years	6 issues, two years	\$5.00 per set
\$20.00	3 years	9 issues, three years	\$7.50 per set
		12 issues, four years	\$9.00 per set

## The Journal of Educational Thought

Faculty of Education, The University of Calgary  
2920 - 24 Ave. N.W., Calgary, Alberta, Canada  
T2N 1N4



## PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by author's last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.



